This Page Is Inserted by IFW Operations and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents will not correct images, please do not report the images to the Image Problem Mailbox.

(19) 日本国特許庁 (JP)

(12) 公開特許公報(A)

(11)特許出願公開番号 特開2000-201688 (P2000-201688A)

(43)公開日 平成12年7月25日(2000.7.25)

(51) Int.Cl.7	截別記号	ΡΙ	テーマコード(参考)
C 1 2 N 15/09	ZNA	C12N 15/00 ZNA.	A 4B024
C 0 7 K 14/72	•	C 0 7 K 14/72	4B063
C 1 2 N 5/10		C 1 2 P 21/02	C 4B064
C 1 2 P 21/02		C 1 2 Q 1/02	4B065
C 1 2 Q 1/02		C 1 2 N 5/00	B 4H045
	審査請求	未請求 請求項の数12 OL (全 23	頁) 最終頁に続く
(21)出願番号	特願平11-98787	(71) 出顧人 000002093	
		住友化学工業株式会社	
(22)出顧日	平成11年4月6日(1999.4.6)	大阪府大阪市中央区北江	兵4丁目5番33号
		(72) 発明者 大下 博文	
(31)優先権主張番号	特願平10-319465	兵庫県宝塚市高司4丁	目2番1号 住化テ
(32)優先日	平成10年11月10日(1998.11.10)	クノサービス株式会社	4
(33)優先権主張国	日本 (JP)	(72) 発明者 徐 明旭	
		兵庫県宝塚市高司4丁	32番1号 住化テ
		クノサービス株式会社内	4
•		(74)代理人 100093285	•
		弁理士 久保山 隆	(外1名)
•			
			最終頁に続く

(54) 【発明の名称】 エストロジェンレセプター遺伝子およびその利用

(57)【要約】

【課題】化学物質のエストロジェン様作用を測定するための方法として、化学物質のエストロジェンレセプター活性化能を評価するための試験系を提供可能とすること。

【解決手段】配列番号1で示されるアミノ酸配列を有する蛋白質をコードするエストロジェンレセプター遺伝子等。

【特許請求の範囲】

【請求項1】配列番号1で示されるアミノ酸配列を有する蛋白質をコードするエストロジェンレセプター遺伝子。

【請求項2】配列番号3で示されるアミノ酸配列からなる蛋白質をコードするエストロジェンレセプター遺伝子。

【請求項3】配列番号2で示される塩基配列を有するエストロジェンレセプター遺伝子。

【請求項4】配列番号4で示される塩基配列からなるエストロジェンレセプター遺伝子。

【請求項5】請求項1~4記載のエストロジェンレセプター遺伝子を含有するベクター。

【請求項6】エストロジェンレセプター遺伝子に宿主細胞で機能可能なプロモーターが機能可能な形で結合されてなる請求項5記載のベクター

【請求項7】請求項1~4記載のエストロジェンレセプター遺伝子が宿主細胞に導入されてなる形質転換体。

【請求項8】請求項5または6記載のベクターが宿主細胞に導入されてなる形質転換体。

【請求項9】宿主細胞が動物細胞である請求項7または 8記載の形質転換体。

【請求項10】請求項7~9記載の形質転換体を培養してエストロジェンレセプターを産生させ、これを回収することを特徴とするエストロジェンレセプターの製造方法。

【請求項11】配列番号1で示されるアミノ酸配列を有するエストロジェンレセプター。

【請求項12】化学物質のエストロジェンレセプター活性化能を評価するためのレボーターアッセイにおいて、エストロジェン応答配列を含む転写制御領域の下流に連結されたレポーター遺伝子と請求項1~4記載のエストロジェンレセプター遺伝子とがエストロジェンレセプター非内在性宿主細胞に導入されてなる形質転換体に、化学物質を作用させることを特徴とする化学物質のエストロジェンレセプター活性化能の評価方法。

【発明の詳細な説明】

[0001]

【発明の属する技術分野】本発明は、エストロジェンレ セプター遺伝子およびその利用に関する。

[0002]

【従来の技術および発明が解決しようとする課題】近年、環境中の幾つかの化学物質がエストロジェン様作用を示すことが報告されている。かかる化学物質の作用はヒトのホルモンバランスを崩し、疾患の原因となることが危惧されることから、化学物質の安全性評価の一環として化学物質のエストロジェン様作用を測定する試みがなされている。エストロジェンの作用機序として、エストロジェンがエストロジェンの標的細胞に存在するエストロジェンレセブターに結合すると、該レセプターは活

性化され、染色体上のエストロジェン応答配列に結合して該配列の下流に在する遺伝子の発現を促進する。そこで、化学物質のエストロジェン様作用を測定するための方法として、化学物質のエストロジェンレセプター活性化能を評価するための試験系の開発が切望されている。 【0003】

【課題を解決するための手段】本発明者らは、かかる状 況の下、鋭意検討した結果、水生動物のモデル動物であ るメダカのエストロジェンレセプターをコードする遺伝 子を見出し、本発明に至った。即ち、本発明は、配列番 号1で示されるアミノ酸配列を有する蛋白質をコードす るエストロジェンレセプター遺伝子(以下、本発明遺伝 子と記す。)、配列番号3で示されるアミノ酸配列から なる蛋白質をコードする該遺伝子、配列番号2で示され る塩基配列を有する該遺伝子、配列番号4で示される塩 基配列からなる該遺伝子、該遺伝子を含有するベクター (以下、本発明ベクターと記す。)、該遺伝子が宿主細 胞に導入されてなる形質転換体(以下、本発明形質転換 体と記す。)、該形質転換体を培養してエストロジェン レセプターを産生させ、これを回収することを特徴とす るエストロジェンレセプターの製造方法、配列番号1で 示されるアミノ酸配列を有するエストロジェンレセプタ ー、化学物質のエストロジェンレセプター活性化能を評 価するためのレポーターアッセイにおいて、エストロジ ェン応答配列を含む転写制御領域の下流に連結されたレ ボーター遺伝子と本発明遺伝子とがエストロジェンレセ ブター非内在性宿主細胞に導入されてなる形質転換体 に、化学物質を作用させることを特徴とする化学物質の エストロジェンレセプター活性化能の評価方法、を提供 するものである。

[0004]

【発明の実施の形態】以下、本発明について詳細に説明 する。本発明遺伝子は、例えば、ヒメダカ等のメダカか ら、J. Sambrook, E. F. Frisch, T. Maniatis著; モレキュラ ー クローニング第2版(Molecular Cloning 2nd editi on)、コールドスプリング ハーバー ラボラトリー (Co 1d Spring Harbor Laboratory) 発行、1989年等に 記載の遺伝子工学的方法に準じて取得することができ る。具体的には、まず、ヒメダカ等のメダカからRNA を調製する。例えば、ヒメダカの内臓を塩酸グアニジン やグアニジンチオシアネート等の強力な蛋白質変性剤を 含む溶液中で粉砕し、さらに該粉砕物にフェノール、ク ロロホルム等を加えることにより蛋白質を変性させる。 変性蛋白質を遠心分離等により除去した後、回収された 可溶性画分から、塩酸グアニジン/フェノール法。SD S-フェノール法、グアニジンチオシアネート/CsC 1法等の方法により全RNAを抽出する。なお、これら の方法に基づいた市販のRNA調製用キットとしては 例えばISOGEN (ニッポンジーン製)がある。得られた全 RNAを鋳型として使用し、該RNAにオリゴdTプラ

イマーをアニールさせた後に逆転写酵素を作用させるこ とにより一本鎖cDNAを合成し、次いで、該一本鎖c DNAに大腸菌RNaseHおよび大腸菌のDNAポリ メラーゼIを作用させて二本鎖のcDNAを合成する。 更に該cDNAの両末端をT4DNAポリメラーゼによ り平滑化する。得られたCDNAはフェノールークロロ ホルム抽出、エタノール沈殿等の通常の方法により精製 し、回収する。なお、これらの方法に基づいた市販ので DNA合成用キットとしては、例えばcDNA合成シス テムプラス (アマシャム社製) がある。このようにして 得られたcDNAを例えば、プラスミドpUC118やファー ジλgt11などのベクターにリガーゼを用いて挿入するこ とによりcDNAライブラリーを作製する。次に、この ようなcDNAライブラリーから、例えば、配列番号2 で示される塩基配列の部分塩基配列を有するDNA断片 をプローブとして用いるハイブリダイゼーション法や、 配列番号2で示される塩基配列の部分塩基配列を有する オリゴヌクレオチドをプライマーとして用いるPCR法 により、本発明遺伝子を取得することができる。また、 上記のようにして調製された全RNAを鋳型に使用して 逆転写反応を行なった後、得られたDNAを鋳型にして PCRを行なうことにより(RT-PCR法)本発明遺 伝子を取得することもできる。上記のPCR法またはR T-PCR法においてPCRにより本発明遺伝子を増幅 する際に用いるプライマーとしては、例えば、20bpから 40bp程度の長さでかつGまたはC塩基の割合が40%から70% 程度の塩基配列を、配列番号2で示される塩基配列の 5、末端領域および3、末端領域からそれぞれ選択し、 該塩基配列を有するオリゴヌクレオチドを合成するとよ い。具体的には、例えば、フォワードプライマーの塩基 配列としては5'-ATG TAC CCT GAA GAG AGC CGG G-3'や 5'-AAG CTTCAT GAG TAA GAG ACA GAG C-3'があげられ、 リバースプライマーの塩基配列としては5°-TCA GTC TTG AAG GGC CGG GGA G-3'があげられる。このようにしてP CRで増幅された本発明遺伝子は、例えば、J. Sambrook. E.F.Frisch, T. Maniatis著; モレキュラー クローニング 第2版 (Molecular Cloning 2nd edition)、コールド スプリング ハーバー ラボラトリー (Cold Spring Harb or Laboratory)発行、1989年等に記載の遺伝子工 学的方法に準じてベクターにクローニングすることがで きる。具体的には例えば、TAクローニングキット(Invi trogen社) やpBluescriptII (Stratagene社) などの市 販のプラスミドベクターを用いてクローニングすること ができる。尚、本発明遺伝子は、配列番号2や配列番号 4で示される塩基配列に基づいて、例えばホスファイト ・トリエステル法 (Hunkapiller, M.et al., Nature, 31 0.105, 1984) 等の通常の方法に準じて、核酸の化学合 成を行うことにより調製することもできる。得られた本 発明遺伝子の塩基配列は、Maxam Gilbert法(例えば、M axam, A.M. &; W. Gilbert, Proc. Natl. Acad. Sci. USA, 74, 5

60, 1977 等に記載される)やSanger法 (例えばSanger, F. &; A.R. Coulson, J.Mol.Biol., 94, 441, 1975、Sanger, F. &; Nicklen and A.R. Coulson., Proc.Natl.Acad.Sci.USA, 74, 5463, 1977等に記載される) に準じて解析することにより確認することができる。

【0005】このようにして取得される本発明遺伝子 を、例えば、宿主細胞内で複製可能なDNAであって、宿 主細胞からの単離、精製が可能であり、検出可能なマー カー遺伝子をもつベクターに、通常の遺伝子工学的手法 を用いて組込むことにより本発明ベクターを構築するこ とができる。本発明ベクターの構築に用いることができ るベクターとしては、具体的には、微生物である大腸菌・ を宿主細胞とする場合、例えば、プラスミドpUC119(宝 酒造 (株) 製) や、ファージミドpBluescriptH (スト ラタジーン社製)等をあげることができ、酵母を宿主細 胞とする場合は、プラスミドpACT2 (Clontech社製)な どをあげることができる。また、哺乳類動物細胞を宿主 細胞とする場合は、pRC/RSV、pRC/CMV (Invitrogen社 製) 等のプラスミド、ウシパピローマウイルスプラスミ ドpBPV(ファルマシア社製)、EBウイルスプラスミドpC EP4 (Invitrogen社製)等のウイルス由来の自律複製起 点を含むベクター、ワクシニアウイルス等のウイルスな どをあげることができ、昆虫類動物細胞(以下、昆虫細 胞と記す。)を宿主細胞とする場合は、バキュロウイル ス等の昆虫ウイルスをあげることができる。バキュロウ イルスやワクシニアウイルス等のウイルスに本発明遺伝 子を組込むには、使用しようとするウイルスのゲノムと 相同な塩基配列を含有するトランスファーベクターを用 いる。このようなトランスファーベクターの具体的例と しては、Pharmingen社から市販されているpVL1392,pVL1 393 (Smith, G.E., Summers M.D. et al.: Mol. Cell. Biol., 3:2156-2165, 1983) pSFB5 (Funahashi, S.et al., J. Vir ol.,65:5584-5588,1991)などのプラスミドをあげること ができる。本発明遺伝子を前記のようなトランスファー ベクターに挿入し、該トランスファーベクターとウイル スゲノムとを同時に宿主細胞に導入すると、トランスフ ァーベクターとウイルスゲノムとの間で相同組換えが起 こり、本発明遺伝子がゲノム上に組み込まれたウイルス を得ることができる。ウイルスゲノムとしては、Baculo virus、Adenovirus、Vacciniavirusなどのゲノムを用い ることができる。本発明遺伝子の上流に、宿主細胞で機 能可能なプロモーターを機能可能な形で結合させ、これ を上述のようなベクターに組み込むことにより、本発明 遺伝子を宿主細胞で発現させることの可能な本発明ベク ター(以下、木発明発現ベクターと記す。)を構築する ことができる。ここで、「機能可能な形で結合させる」 とは、本発明遺伝子が導入される宿主細胞においてプロ モーターの制御下に発現するように、該プロモーターと 本発明遺伝子とを結合させることを意味する。使用する プロモーターは、形質転換する宿主細胞内でプロモータ

一活性を示すものであれば特に制限はなく、例えば、宿 主細胞が動物細胞や分裂酵母である場合は、例えば、ラ ウス肉腫ウイルス (RSV) プロモーター、サイトメガロ ウイルス (CMV) プロモーター、シミアンウイルス (SV4 0)の初期もしくは後期プロモーター、マウス乳頭腫ウ イルス (MMTV) プロモーター、単純ヘルベスウイルス (HSV) のチミジンキナーゼ(tk) 遺伝子プロモーター 等をあげることができる。宿主細胞が出芽酵母である場 合はADH1プロモーターなどをあげることができる。ま た、宿主細胞において機能するプロモーターをあらかじ め保有するベクターを使用する場合は、ベクター保有の プロモーターと本発明遺伝子とが機能可能な形で結合す るように、該プロモーターの下流に本発明遺伝子を挿入 すればよい。例えば、前述のブラスミドpRC/RSV,pRC/CM V等は、動物細胞で機能可能なプロモーターの下流にク ローニング部位が設けられており、該クローニング部位 に本発明遺伝子を挿入し動物細胞へ導入すれば、本発明 遺伝子が発現する。これらのプラスミドにはあらかじめ SV40の自律複製起点(ori)が組み込まれているため、o ri(-)のSV40ゲノムで形質転換された培養細胞、例えばC OS細胞等に該プラスミドを導入すると、細胞内でプラス ミドのコピー数が非常に増大し、結果として該プラスミ ドに組み込まれた本発明遺伝子を大量発現させることも できる。また、前述の酵母用プラスミドpACT2はADIIIプ ロモーターを有しており、該プラスミドまたはその誘導 体のADH1プロモーターの下流に本発明遺伝子を挿入すれ ば、本発明遺伝子を例えばCG1945 (Cloutech社製)等の 出芽酵母内で大量発現させることが可能な本発明ベクタ ーが構築できる。

【0006】上述のようにして構築された本発明ベクタ ーを宿主細胞に導入することにより、本発明形質転換体 を取得することができる。本発明ベクターを宿主細胞へ 導入する方法は、形質転換される宿主細胞に応じて通常 用いられる方法でよい。例えば、大腸菌を宿主細胞とす る場合は、「モレキュラー・クローニング」(J. Sambrook ら、コールド・スプリング・ハーバー、1989年)等 に記載される塩化カルシウム法やエレクトロボレーショ ン法等を用いることができ、酵母菌を宿主細胞とする場 合は、例えばリチウム法に基づくYeast transformation kit(Clontech社製)などを用いてベクターを導入するこ とができる。また、哺乳類動物細胞や昆虫細胞等の動物 細胞を宿主細胞とする場合は、例えば、リン酸カルシウ ム法、DEAEデキストラン法、エレクトロボレーション 法、またはリポフェクション法等により該宿主細胞に本 発明ベクターを導入することができる。尚、ウイルスを ベクターに用いる場合は、上述のような一般的な遺伝子 導入法によりウイルスゲノムを宿主細胞に導入できるほ か、ウイルスゲノムを含有するウイルス粒子を宿主細胞 へ感染させることによってもウイルスゲノムを宿主細胞 に導入することができる。

【0007】本発明形質転換体の選抜は、導入された本 発明ベクターが有する検出マーカー遺伝子の性質に応じ た方法を用いればよい。例えば、検出マーカー遺伝子 が、細胞致死活性を示す薬剤に対する耐性遺伝子である 場合には、該薬剤を添加した培地を用いて、本発明ベク ターを導入した細胞を培養すればよい。このようにして 用いることのできる薬剤耐性遺伝子と選抜薬剤との組み 合わせとしては、例えば、ネオマイシン耐性遺伝子とネ オマイシンとの組合せ、ハイグロマイシン耐性遺伝子と ハイグロマイシンとの組み合せ、プラストサイジンS耐 性遺伝子とブラストサイジンSとの組合せ等をあげるこ とができる。また、検出マーカー遺伝子が、宿主細胞の 栄養要求性を相補する遺伝子である場合には、該栄養素 を含まない最少培地を用いて、本発明ベクターを導入し た細胞を培養すればよい。さらに、本発明発現ベクター を導入した場合は、エストロジェン結合活性に基づく検 出方法を用いることもできる。本発明遺伝子が宿主細胞 の染色体に導入されてなる本発明形質転換体を取得する には、例えば、本発明ベクターを制限酵素等で消化する ことにより直鎖上にした後、これを前述の方法で宿主細 胞へ導入して該細胞を通常数週間培養し、導入された本 発明ベクターにコードされる検出マーカーを指標にして 目的とする形質転換体を選抜すればよい。例えば、上記 のような選択薬剤に対する耐性遺伝子を検出マーカー遺 伝子として持つ本発明ベクターを前述の方法で宿主細胞 に導入し、選択薬剤を添加した培地で数週間以上該細胞 を継代培養して、コロニー状に生き残った選択薬剤耐性 クローンをピペットで吸い上げ純化することにより、本 発明遺伝子が宿主細胞の染色体に導入されてなる本発明 形質転換体を取得することができる。該形質転換体は、 凍結保存が可能であり必要に応じて起眠して使用するこ とできるので、一過性の遺伝子導入株と比較して、形質 転換体作製の手間を省くことができ、形質転換体の性能 を一定に保つこともできる。

【0008】上述のようにして得られた本発明形質転換 体を培養することにより本発明のエストロジェンレセプ ターを産生させることができる。例えば、本発明形質転 換体が微生物である場合、該形質転換体は、一般微生物 における通常の培養に使用される炭素源や窒素源、有機 ないし無機塩等を適宜含む各種の培地を用いて培養すれ ばよい。培養は、一般微生物における通常の方法に準じ て行い、固体培養、液体培養(試験管振とう式培養、往 復式振とう培養、ジャーファーメンター (Jar Fermente r) 培養、タンク培養等)等が可能である。培養温度 は、微生物が生育する範囲で適宜変更でき、例えば、約 15℃~約40℃の培養温度、約6~約8の培地別で培 養するとよい。培養時間は、種々の培養条件によって異 なるが、通常約1~約5日間である。また、上記形質転 換体が動物細胞である場合、一般の培養細胞における通 常の培養に使用される培地を用いて培養すればよい。選

択薬剤を利用して当該形質転換体を選抜した場合は、対 応する選択薬剤を共存させて培養するのが望ましい。哺 乳類動物細胞の場合、例えば10v/v%となるようFBSを添 加したDMEM培地等の培地を用いて、37℃、5v/v%CO2 存在下等にて、培地を数目ごとに交換しながら培養す る。細胞がコンフルエントになるまで増殖したら、0. 25w/v%程度のトリプシンPBS溶液を用いて個々の細胞 に分散させ、数倍に希釈して新しい培養容器に播種し培 養を続ける。目的とする量まで細胞が増殖したら細胞を 集める。昆虫細胞の場合も同様に、10v/v%FBSおよび2w/ v%Yeastlateを含むGrace's medium等の昆虫細胞用培地 を用いて25℃から35℃で継代培養する。ただし、Sf 21細胞などの培養容器からはがれやすい細胞の場合は、 トリプシン液ではなくピペッテイングにより細胞を分散 させ継代を行なう。また、Baculovirus等のウイルスへ クターを含む形質転換体の場合は、細胞質効果により細 胞が死滅する前、例えば培養開始から72時間目までに 培養を終了することが好ましい。本発明形質転換体によ り産生されたエストロジェンレセプターの回収は、適 宜、通常の単離、精製の方法を組み合わせて行えば良 く、例えば、培養終了後、形質転換体の細胞を遠心分離 等で集め、該細胞を通常のバッファー、例えば、20mMHE PES pH7,1mM EDTA,1mM DTT,0.5mM PMSFからなるバッフ ァー等に懸濁した後、ポリトロン、超音波、ダウンスポ モジナイザー等を用いて細胞を破砕し、破砕液を数万x gで数十分から1時間程度超遠心分離し、上清画分を回 収することにより、エストロジェンレセプターを含む画 分を得ることができる。さらに、前記上清画分をイオン 交換、疎水、ゲルろ過、アフィニティ等の各種クロマト グラフィーに供することにより、より精製されたエスト ロジェンレセプターを回収することもできる。この際、 後述のエストロジェンレセプターが結合する塩基配列を 含む156pから2006p程度の長さのオリゴヌクレ オチドをプローブとしたDNA結合アッセイなどにより、 目的とするエストロジェンレセプターを含む画分を見分 けることができる。このようにして製造された本発明の エストロジェンレセプターは、例えば、エストロジェン レセプターに対する化学物質の親和性を測定するための ラジオレセプターアッセイ等に用いることができる。 【0009】上述のようにして構築された本発明発現べ クターは、例えば、化学物質のエストロジェンレセプタ 一活性化能を評価するためのレポーターアッセイに利用 することができる。具体的には、エストロジェン応答配 列を有しエストロジェンレセプターにより転写が制御さ れる遺伝子、例えばビテロジェニン遺伝子の転写制御領 域の下流にレポーター遺伝子を結合させたキメラ遺伝 子、または、エストロジェン応答配列の下流に転写開始 に必要な塩基配列とレポーター遺伝子とを結合させたキ メラ遺伝子(以下、本キメラ遺伝子と記す。)を、細胞 内でのエストロジェンレセプターの転写調節能をモニタ

ーするためのレポーター遺伝子として用いる。エストロ ジェン応答配列 (estrogen response element)とは、 エストロジェンにより転写が制御される遺伝子のプロモ ーターの上流に存在し、エストロジェンレセプターによ って認識される塩基配列を意味する。エストロジェンの 結合したエストロジェンレセプターは活性化されてエス トロジェン応答配列に結合することにより、該配列の下 流にある遺伝子の転写を促進する。エストロジェン応答 配列のコンセンサス配列としては、塩基配列 AGGTCAXXX TGACCT (Xは、A、G、C、またはTを意味する。) が一般 に知られている。レポーター遺伝子としては、ルシフェ ラーゼ遺伝子、分泌型アルカリフォスファターゼ遺伝 子、βガラクトシダーゼ遺伝子、クロラムフェニコール アセチルトランスフェラーゼ遺伝子、成長ホルモン遺伝 子などが利用できる。上述のように作製した本発明発現 ベクターと、本キメラ遺伝子を組み込んだベクターと を、内在性のエストロジェンレセプターを産生していな い宿主細胞、例えばIfeLa細胞やNIH3T3細胞などに導入し 形質転換体を取得する。この形質転換体をそのまま1日 から数日間培養する間に、例えばエストロジェン様作用 をもつ化学物質を培地中に加えて前記形質転換体に作用 させる。該形質転換体が産生するエストロジェンレセプ ターが化学物質の結合により活性化された場合は、レポ ーター遺伝子のmRNAへの転写が促進され、ルシフェラー ゼ酵素蛋白質が形質転換体の細胞内に蓄積する。この状 態の形質転換体を破砕して細胞粗抽出物を調製し、レポ ーターの酵素活性等を指標にして細胞当たりのレポータ 一蛋白質の量を求める。例えば、レポーター遺伝子とし てルシフェラーゼ遺伝子を用いた場合、前記細胞租抽出 物にルシフェラーゼの基質であるルシフェリンを加える と発光し、発光量はルシフェラーゼ量に比例する。従っ て、この発光量をルミノメーター等の測定装置で測定す ることにより、ルシフェラーゼ遺伝子の発現量がわか り、よって、その際に添加されていた化学物質のエスト ロジェンレセプター活性化能を評価することができる。 また、本発明発現ベクターと、本キメラ遺伝子が組み込 まれたベクターとを同時に宿主細胞に導入して、本発明 遺伝子および本キメラ遺伝子が宿主細胞の染色体に導入 されてなる形質転換体を取得し、上記レポーターアッセ イに用いてもよい。該形質転換体は凍結保存が可能であ り必要に応じて起眠して使用することできるので、これ を一旦取得すると、アッセイの度ごとにこれらの遺伝子 を宿主細胞に導入して新たな形質転換体を取得する必要 が無く、また、形質転換体の性能を一定に保つこともで きることから、例えばハイスループットスクリーニング 等の自動化された大規模スクリーニングを実施する際に 有用である。

[0010]

【実施例】以下、実施例により本発明を更に詳細に説明 するが、本発明はこれら実施例によって限定されるもの

0℃ 1-1.5min、72℃ 1-3min)を行った。

ではない。

【0011】実施例1(本発明遺伝子の取得)

餌(コイ稚魚用) にβーエストラジオール(和光純薬工業株式会社製)10mg/1を10mg/g餌となるように添加し、これにアセトンを加えてよく混和した後、ヘアードライヤーの下でアセトンを除去した。こうして得た処理餌を、約3ヶ月のヒメダカ雌10個体に3回/日の頻度で1日間飽食量与えた。βーエストラジオール投与24時間後に、これらのヒメダカから内臓を摘出し、直ちに組織1

XU1 : 5'-ATG TAC CCT GAA GAG AGC CGG G-3' (22mer, GC/AT = 13/9)

XU14: 5'-TCA GTC TTG AAG GGC CGG GGA G-3' (22mer, GC/AT = 14/8)

次いで、前記RT-PCRで得られたDNA断片をpCR2.1 (TA cloningベクター) にサブクローニングして大腸菌DH5-α に導入し、プラスミドを調製した(pCR-ER)。pCR2.1の塩基配列に基づくプライマーおよび上記のプライマー (XU1, XU14)を用い、 ABI sequence systemで、前記のpCR-ERにクローニングされたDNA断片の塩基配列を決定した。その結果、配列番号2で示される塩基配列が明らか

となった。また、上記のようにして調製したRNA1μ8を 鋳型とし、ランダム9merプライマーを用い、子め30℃で 10分間逆転写反応を行い、引き続き42℃で50分間逆転写 反応を行った。続いて下記のプライマー(XU36とXU14) を用い PCR(30サイクル、94℃ 30sec-1mim、50-60℃ 1 -1.5min、72℃ 1-3min)を行った。

gあたり10mlのトリゾール試薬を加えてホモジナイズし

た後、クロロホルムを加えて遠心分離した。水相を採取

してイソプロパノールを加えRNAを沈澱させた。約0.3g

約500μgのRNAが得られた。このようにして調製したRNA

1μgを鋳型とし、ランダム9merプライマーを用い、子め

30℃で10分間逆転写反応を行い、引き続き42℃で50分間

逆転写反応を行った。続いて下記のプライマー (XU1とX

U14) を用い PCR (30サイクル、94°C 30sec-1min、50-6

XU36: 5'-AAG CTT CAT GAG TAA GAG ACA GAG C-3'(25mer GC/AT = 11/14)

XU14: 5'-TCA GTC TTG AAG GGC CGG GGA G-3' (22mer, GC/AT = 14/8)

次いで、前記RT-PCRで得られたDNA断片をpCR2.1 (TA cloningベクター) にサブクローニングして大腸菌DH5- α に導入し、プラスミドを調製した(pCR-ER2)。pCR2.1の 塩基配列に基づくプライマーおよび上記のプライマー (XU36, XU14)を用い、 ABI sequence systemで、前記の pCR-ER2にクローニングされたDNA断片の塩基配列を決定した。その結果、配列番号4で示される塩基配列が明らかとなった。

【 0 0 1 2 】実施例2(本発明遺伝子発現用ベクターの 構築)

実施例1で得られたプラスミドpCR-ERからエストロジェンレセプター遺伝子をXba I とllind IIIで切り出し、同じ制限酵素で消化した発現ベクターpRc/RSVに組みこみ、エストロジェンレセプターを発現させるための発現

ベクターRSV-ERを構築した。具体的な過程を図1に、発現ベクターRSV-ERの構造の詳細を図2示す。また、同様にして、実施例1で得られたプラスミドpCR-ER2からエストロジェンレセプター遺伝子をXba I とHind IIIで切り出し、同じ制限酵素で消化した発現ベクターpRc/RSVに組みこみ、エストロジェンレセプターを発現させるための発現ベクターRSV-ER2を構築した。

【0013】実施例3(エストロジェンレセプターに応答するレポーター遺伝子を有するベクターの構築) 既知のセノプスのA2ビテロジェニン遺伝子(GenBank Ac cession No. X00205)の5、末端領域のエストロジェン 応答配列(以下、EREと記す。)のコンセンサス配列を もとに、下記のオリゴヌクレオチド1およびオリゴヌク レオチド2を合成した。

オリゴヌクレオチド1:

5'-CCA AAG TCA GGT CAC AGT GAC CTG ATC AAA GGA AC-3'オリゴヌクレオチド2:

5'-CTT TGA TCA GGT CAC TGT GAC CTG ACT TTG GGT TC-3

両オリゴヌクレオチドの末端をカイネーションによりリン酸化した。カイネーション反応液は、10 nmolのオリゴヌクレオチド1または10 nmolのオリゴヌクレオチド、2、5μ1の10 Xカイネーション バッファー、1μ1の10 mMのATP、2.5μ1のポリメクレオチドキナーゼ(宝酒造社製)を1.5 ml容チューブに採り、減菌蒸留水を加え全量を50μ1として調製した。カイネーション反応は37℃で1時間行った。反応終了後、リン酸化したオリゴヌクレオチド1とオリゴヌクレオチド2をアニーリングさせ、2本鎖のEREを得た。アニーリング反応液は、リン酸化させたオリゴヌクレオチド1およびオリゴヌクレオチド2のそれぞれ20μ1ずつを、1.5 ml容チュ

ーブに加え、95℃で5分間保温し、60℃、次いで37℃に てそれぞれ1時間保温した後、室温で約1時間放置し た。反応終了後、10µ1のアニーリング反応液に10µ1の DNAリガーゼ(ライゲーションキット、宝酒造社製)を 加え、2本鎮のERE断片を連結した。反応液をアガロー スゲル電気泳動に供してDNA断片の長さを分析し、ER E断片が4個連結されたと判断されるDNA断片(以下、 4 X ERE断片と記す。)およびERE断片が5個連結された と判断されるDNA断片(以下、5 X ERE断片と記 す。)をそれぞれ回収した。これらのDNA断片をブランティングキット(宝酒造社製)を用い、末端を平滑化 した。一方、pBluescript(SK-)を制限酵素EcoR I(宝酒

造社製)で切断し、5'末端をアルカリフォスフェターゼ (宝酒造社製)で脱リン酸化した。前記のDNA断片(4 X ERE断片または5 X ERE断片)とpBluescript(SK-)とを それぞれDNAライゲーションキットを用いて結合した。 得られた反応液で大腸菌DH5αのコンピテントセル (東洋紡社製)を形質転換してアンピシリン耐性となっ た株を選抜し、該アンピシリン耐性株からプラスミドDN Aを調製し、塩基配列を ABI PRISMT*377 DNA Sequence System(パーキンエルマージャパン社製)で確認した。次 に、 HSV tKプロモーター配列を持つベクターpTKB (ク ロンテック)を制限酵素Sal IおよびXho I (それぞれ宝 酒造社とニッポンジーン社製)で切断した後アガロース ゲル電気泳動で分析し、約1 kbpのtKプロモーター断片 を得た。該断片をブランティングキットを用い、末端を 平滑化した。一方、ルシフェラーゼ遺伝子を持つレポー タープラスミド pGL-3 (ピッカジーン)を制限酵素Sma I (宝酒造社製)で切断し、5 末端をアルカリフォスフ ェターゼ (宝酒造社製)で脱リン酸化した。2つのDNA 断片をDNAライゲーションキットを用いて結合した。得 られた反応液で大腸菌DH5αのコンピテントセル(東 洋紡)を形質転換してアンピシリン耐性となった株を選 抜し、該アンピシリン耐性株からブラスミドDNAを調製 し、 tkプロモーター挿入したレポータープラスミド(tk -pGL-3)を取得した。次に、上記の4 X ERE断片が挿入さ れたpBluescriptを制限酵素KpnIおよびXbaI(それぞれ 宝酒造社とニッポンジーン社製)で切断し、アガロース ゲル電気泳動で4 X ERE断片を回収した。一方、tK-pGL-3をKpn 1およびNhe I (いずれも宝酒造社製)で消化 し、5 末端をアルカリフォスフェターゼ(宝酒造社 製)で脱リン酸化した。このようにして調製された4 X ERE断片とtK-pGL-3とをDNAライゲーションキットを用い て結合した。得られた反応液で大腸菌DH5α株のコン ピテントセル (東洋紡)を形質転換してアンピシリン耐 性となった株を選抜し、該アンビシリン耐性株からプラ スミドDNAを調製し、4 X ERE 断片およびtKプロモータ 一断片を保有するレポータープラスミド(ERE-tK-pGL)を 取得した。該プラスミドの構築の過程を図るに、該プラ スミドの構造を図4に示す。また、同様にして、上記の 5 X ERE断片が挿入されたpBluescriptから5 X ERE 断 片を回収し、該断片と、Kpn TおよびNhe Tで消化された tK-pGL-3とを結合させ、5 X ERE 断片およびtKプロモ ーター断片を保有するレポータープラスミド(ERE5-tK-p GL)を取得した。

【0014】実施例4(プラスミドDNAの大量調製) 実施例2で得られた発現ベクターRSV-ER、レポータープ ラスミドERE-tK-pGLおよびERE5-tK-pGL、ならびにコン トロールレポータープラスミドであるpRL-tK(ビッカジーン)のDNAを以下の方法より大量に調製した。上記プラスミドを含む大腸菌をアンプシリン(終濃度50μg/ml)を含有LB培地3mlに植菌し、37℃で一晩振動培養した。そ の培養液をアンピシリン (50μg/ml)を含むLB培地200ml に植菌し、一晩振動培養した。一晩培養後の菌体を5,00 0 rpm, 10 分間, 4℃で遠心分離し(CR21、日立工 機)、得られた沈澱を0.1 M のSTEバッファー60 mlに懸 濁し、同条件で再度遠心分離した。沈澱を3 ml のsolut ion 1に懸濁し、1 ml のリゾチーム(20 mg/ml)を加え、 室温で5分間放置した。引き続き10 mlの Solution 2を 加え、氷上に10分間放置した後、7.5 ml のsolution 3 を加え、氷上に15分間放置した。12,000 rpm, 20分, 4 ℃で遠心分離し、上清を50 mlチューブに移し、0.6容量 のイソプロピルアルコールを加え、室温で15分間放置し た。3,000 rpm, 10分間,室温で遠心分離し(CR5DL、日 立工機)、70% エタノールで洗浄し、乾燥させた。沈澱 を4.2 mlの TE バッファーに溶かし、5 mg/ml Rnase溶 液(ニッポンジーン)を28µ1加え、50℃,30分間インキ ュベートした。2 mg/mlエチジウムブロマイド溶液400μ 1とCsCl (関東化学) 4.6 g を加えた後、日立工機シール チューブに移し、55,000 rpm,20℃,16時間遠心分離(S CP85H2, 日立工機) した。スーパーコイルドブラスミドDN Aのバンドを注射筒で抜き取り、55,000 rpm, 20℃. 16 時間再度遠心分離した。再びスーパーコイルプラスミド DNAのバンドを注射筒で抜き取り、水飽和イソアミルア ルコールでエチジウムプロマイドを完全に除き、一晩5 mM STEバッファーに透析した後、試料として用いた。 【0015】実施例5(細胞の培養)

不活化済み牛胎仔血清 (GIBCO-BRL、米国) を活性炭ー デキストランで処理し、細胞培養の培地作製に用いた。 処理過程でおける各ステップは以下の通りであった。25 Mスクロース(和光純薬)、1.5mM MgCl₂(和光純薬)、 10mM HEPES (pH7.4) (同仁化学、熊本) 1L中にノーリッ トEXW (ナカライテスク) 2.5gとデキストランT-70 (フ ァルマシアバイオテク、スウェーデン)0.25gを懸濁 し、4°Cで終夜攪拌した。本懸濁液を12,000rpmで10分遠 心 (CR21、日立工機)して活性炭を沈殿させた。これ を不活化済み牛胎仔血清 (GIBOD-BRL、米国) 1Lに懸濁 し、4°Cで終夜撹拌した。その後、12,000rpmで10分遠心 (CR21、日立工機)して活性炭を沈殿させ、取り除い た。以上の操作を2回繰り返した後、ザルトラブV500 (0.20gm、ザルトリウス、独国)を用いてフィルター **ろ過したろ液を活性炭―デキストラン処理済みの牛胎仔** 血清とした。ヒト子宮癌細胞株Hela(大日本製菜製) を、10%活性炭ーデキストラン処理済みの牛胎仔血清、 0.03% Lーグルタミン (日水製薬) および0.15% 炭酸水 素ナトリウムを添加したイーグルMEM(日水製薬)培地 を用いて10 cmの組織培養用ディッシュ(ファルコン) に培養した。細胞の維代・播種は培地を除去後、適量の PBS(-) (日水製薬)で接着した細胞を洗浄し、 PBS(-)8 Omlに覧トリプシン(DIFCO, 米国) 10ml、0.2% EDTA・3 Na(同仁化学) 10mlを添加した液を用いて細胞を剥離し た。細胞の培養はすべて5% CO₂および飽和湿度下、37°C

でCO2インキュベータ(アステック)内で培養した。 【0016】実施例6(レポーターアッセイ) 以下の操作を各条件ごとに4連で実験を行った。実施例 5で培養した細胞の培地を除去し、PBS(-)で1回洗浄し た。5%トリプシン (DIFCO, 米国) 5 mlを加え、細胞を剥 離した。細胞の計数した後24ウェルマルチウェルプレー ト (ファルコン) に播種した。細胞を1ウェルあたり4 0,000個のHeLa細胞を播種した。1 ウェルあたり0.5mlの 10% 活性炭ーデキストラン処理済みの牛胎仔血清含有培 地を添加し翌日まで培養した。トランスフェクションは 1ウェルあたり0.25~1.0μgのレセプター発現ベクター RSV-ER、0.1~0.5μgのレポータープラスミドERE-tK-pG LまたはERE5-tK-pGL、および、0.05~0.1µgのコントロ ールレポータープラスミドpRL-tKを、0.35μ1/ウェルの リポフェクチン(GIBCO-BRL)またはリポフェクタミン(GI BOD-BRL)を用いて、無血清培地中にて細胞に導入した。 1ウェルあたりの培地量は200μ1とした。導入方法は添 付説明書に従った。各細胞をトランスフェクション培地 中で5時間培養後に 10% 活性炭ーデキストラン処理済 みの牛胎仔血清含有培地に交換し、翌日まで培養を続け た。次いで、Bーエストラジオール(和光純薬社製)を DMSO (関東化学)に溶解しイーグルMEM培地に溶かし て、上述の細胞に添加した。βーエストラジオール終濃 度は10 pMから100 nMの範囲とした。βーエストラジオ ール添加後の1ウェルあたりの培地量は1 mlとした。β ーエストラジオールを添加してから約28時間培養後、培 地を除去し、βーエストラジオールに曝露した細胞をPB S(-)で2回洗浄した。ピッカジーンデュアルキット(東 洋インキ)中の細胞溶解剤を超純水で与倍希釈したもの を1ウェルあたり50μ1添加し、室温で30分間放置して 細胞を溶解した。方法はキットの添付説明書に従った。 細胞溶解液10μ1を白色96ウェルマルチウェルプレ ート(燐光測定用、ベルトールド、ドイツ)に移し、キ ット中の発光基質2種(ルシフェリン、セランテラジ ン)を順次添加し、それぞれの発光量をルミノメーター (ベルトールド、LB9GP)で測定した。はじめに添加す る発光基質はレポータープラスミドERE-tK-pGL由来のホ タルルシフェラーゼによる発光を測定するために用い、 後で添加する発光基質はコントロールレポータープラス ミドpRL-tK由来のシーパンジールシフェラーゼによる発 光を測定するために用いた。前者の発光量は化学物質が 遺伝子の転写活性にあたえる影響を反映している。ま

【 O O 1 7 】実施例7(トランスフェクション条件の検 討)

た、後者の発光量は内部標準として個々の実験データを

実施例6記載のレポーターアッセイにおいて、トランスフェクションの際の各ウェルへのプラスミドの添加量と、リポフェクチンまたはリポフェクタミンの添加量について検討した。まず、各ウェルに、プラスミド(tk-p

GL-3)を0.2μgから1.2μg、リポフェクチンまたはリポ フェクタミンを0.2μ1から0.8μ1の範囲で添加してトラ ンスフェクションを行い、得られた細胞を培養して、実 施例6記載の方法に準じてルシフェラーゼ活性を測定し た。結果を図5に示した。リポフェクチンを用いた場合 には、リポフェクチン0.4µ1とプラスミド0.4µgを加え たときに、リポフェクタミンを用いた場合は、リポフェ クタミンを0.6μ1とプラスミド0.4μg加えたときに、そ れぞれ最も高いルシフェラーゼ活性が認められた。次 に、リホフェクチンまたはリポフェクタミン添加量につ いて更に詳細に検討した。すなわち、各ウェルに加える プラスミド量(0.4μg)を固定し、リポフェクチンはウ ェル当り0.3~0.55以1の範囲で、とリポフェクタミンは ウェル当り0.4~0.9μ1の範囲で添加し、ルシフェラー ゼ活性を測定した。その結果を図6に示した。リポフェ クチンを0.35μ1加えた場合、または、リポフェクタミ ンを0.6 μ1 加えた場合にそれぞれ最も高いルシフェラー ゼ活性を認めた。この結果を基に、リポフェクチンとリ ポフェクタミンの添加量をそれぞれ0.35μ1と0.6μ1 に 固定して、再度プラスミドの添加量を詳細に検討した。 その結果を図7に示した。プラスミドを0.4~0.6μg を 加えた場合に最も高いルシフェラーゼ活性が認められ た。以上の結果を基に、レポーターアッセイにおける各 ウェルに添加する最適なプラスミドの量を0.4µg ~ 0. 6μg、リポフェクチンの量を 0.35μl、リポフェクタミ ンの量を0.6μ1と決定した。

【0018】実施例8 (β -エストラジオールによる 反応性の確認)

実施例6記載の方法および実施例7で得られた最適条件下に、HeLa細胞における β -エストラジオール(E_2)のエストロジェンレセプター活性化能を測定した。ウェル内の β -エストラジオールの終濃度が10pM から50 μ M となる条件でルシフェラーゼ活性を測定した。結果を図8と9に示した。100pM以上でルシフェラーゼ活性の上昇を認め、1 μ 0mでの活性は最大に達した。 μ 0mでは活性の低下が認められたがこれは細胞毒性によるものであると思われた。

【0019】実施例9(化学物質のエストロジュン様作用の測定)

実施例6の方法および実施例7の結果で得られた最適条件下に、HeLa細胞におけるビスフェノールA、pーノニルフェノール、酢酸トリブチルすず、またはフタル酸ジー2ーエチルヘキシルのエストロジェンレセプター活性化能を測定した。アッセイにおける各被験化学物質の終濃度は10 pMから50μMの範囲として実験を行った。各被験化学物質の溶解性を光学顕微鏡下、沈殿物あるいは浮遊物がないことを目視で確認した。また、被験化合物に替えてβーエストラジオールを終濃度500pMとなるように添加した群を設定し、陽性対照とした。ビスフェノールAまたはノニルフェノールでは100nMで活性が上昇し

補正するために用いる。

始め、10μMでは活性化倍率は溶媒コントロール (DMSOのみ添加)の約5倍に上昇した (図10、11)。酢酸トリブチルすずまたはフタル酸ジー2ーエチルヘキシルでは活性の上昇は認められなかった (図12、13)。さらに上記4種について50μMでの実験も行った。その結果を図14に示した。ビスフェノールAで更なる活性の上昇を認めた。ノニルフェノールにおいては細胞に対する毒性に起因すると思われる活性の低下が認められた

【0020】参考例1 (他の細胞用いたアッセイ) 他の種類の培養細胞を選び、実施例5~9と同様の実験 を行う。

[0021]

【発明の効果】メダカエストロジェンレセプターをコードする遺伝子、および、化学物質の該レセプター活性化能を評価するための試験系等が提供可能となる。

【0022】 【配列表】

<;110>; Sumitomo Chemical Company Limited

<;120>; Estrogen receptor genes

<;130>; P150237

<;150>; JP 10/319465 <;151>; 1998-11-10

<;160>; 4

<;210>; 1 <;211>; 575

<;212>; PRT

<:213>: Oryzias lapites

<;400>: 1

Met Tyr Pro Glu Glu Ser Arg Gly Ser Gly Gly Val Ala Ala Val Asp

5 10

Leu Leu Glu Gly Thr Tyr Asp Tyr Ala Ala Pro Asn Pro Ala Thr Thr

Pro Leu Tyr Ser Gln Ser Ser Thr Gly Tyr Tyr Ser Ala Pro Leu Glu
35 40 45

Thr Asn Gly Pro Pro Ser Glu Gly Ser Leu Gln Ser Leu Gly Ser Gly

Pro Thr Ser Pro Leu Val Phe Val Pro Ser Ser Pro Ars Leu Ser Pro 65 70 75 80

Phe Met His Pro Pro Ser His His Tyr Leu Glu Thr Thr Ser Thr Pro

Val Tyr Arg Ser Ser Ilis Gln Gly Ala Ser Arg Glu Asp Gln Cys Gly 100 105 110

Ser Arg Glu Asp Thr Cys Ser Leu Gly Glu Leu Gly Ala Gly Ala Gly 115 120 125

Ala Gly Gly Phe Glu Met Ala Lys Asp Thr Arg Phe Cys Ala Val Cys

130 135 140

Ser Asp Tyr Ala Ser Gly Tyr His Tyr Gly Val Trp Ser Cys Glu Gly 145 150 155 160

Cys Lys Ala Phe Phe Lys Arg Ser IIe Gln Gly His Asn Asp Tyr Met. 165 170 175

Cys Pro Ala Thr Asn Gln Cys Thr 11e Asp Arg Asn Arg Arg Lys Gly 180 185 190

Cy	s Gl	n Ál	la Cy	s Ar	g Lei	u Arg	g Lys	s Cys	s Tyi	r G)ս	ı Val	Gly	Met	Met	Lys
		119	95				200)				205	5		
Gl	y G1,	y Va	ıl Ar	g Ly	s Asj	? Arg	g He	e Arg	ş 11 e	e Lei	ı Arş	g Arg	. Asp	Lys	Arg
	.21	0 -	, .			215	5			*	220)	•		
Ara	g Th	r Gl	y Va	1 G1;	y Asi	• G15	/ Asp	Lys	. Val	l Val	Lys	Gly	Gln	Glu	His
225					230				•	235					240
Lys	s Th	r Va	l Hi	s Ty		Gly	/ Arg	Lys	: Arg	s Ser	Ser	Thr	Gly	Gly	Gly
	*			245					250					255	
Gly	y G1;	y G1		y Gly	y G13	' Arg	; Leu			Thr	Ser	He			Glu
			26					265		_ +		•	270		
Glr	ı Val			u Lei	ı Leu	Gln			Glı	ı Pro	Pro			Cys	Ser
		27		-		_	280					285			
Arg			s Le	u Sei	r Ang			Thr	Glu	ı Val			Met	Thr	Leu
	290		. 14			295					300			_	
		· Se		t Ala				Leu	Val			He	Ala	Trp	
305				. Cl.) , 			C	315		4	C1	0.1	320
Lys	Lys	Le	u rr	o Gly 325		Leu	GIN	Leu			HIS	ASP	GIN		
ريم ا			ر د	r Ser		سم ا	Clu	Va1	330		Ha	C10	1	335	
LCu	L.C.		340		пр	LC.U	Olu	345		nec	116	uly	350	116	пр
Arg	Ser	- 114		s Cys	Pro	Glv	lve			Phe	Δla	G1n		Leu	ملآ
		355		, c, c		013	360		110	1110	mu	365	пор	LCu	110
Leu	Asp			ı Glu	Glv	Asp			Glu	Glv	Met.		Glii	He	Phe
	370					375					380		.,		
Asp			ı Let	ı Ala	Thr			Arg	Phe	Arg		Leu	Lys	Leu	Lys
385					390					395					400
Pro	G] u	Glu	ı Phe	2 Val	Cys	Leu	Lys	Ala	He	He	Leu	Leu	Asn	Ser	
				405					410					415	
Ala	Phe	Ser	Phe	Cys			Thr	Met	Glu	Pro	Leu	llis	Asn	Ser	Ala
			420					425					430		
Ala	Val	Glr	. Ser	Met			Thr	He	Thr	Asp	Ala	Leu	He	His	Tyr
		435		٠.			440					445		•	
				Gly									Arg	Gln	Ala
	450														
	Pro	Leu	Leu	Leu		Ser	His	He	Arg		Met	Ser	Asn	Lys	Gly
465				_	470					475					480
Met.	Glu	His	Leu	Tyr	Ser	Met	Lys	Cys		Asn	Lys	Val	Pro		Tyr
	,			485					490					495	
Asp	Leu	Leu		Glu					llis	Arg	Leu	His		Pro	Val
	.,		500					505					510	_	
Arg	Ala			Ser	Leu	Ser		Val	Asp	Arg	Asp		Pro	Ser	Thr
C	C	515		C1	C1.		520	n	C1	~		525		_	
ser		шу	ыу	Gly	GIY		Ala	ľľo	ыу	Ser		Ser	Ala	Ser	Arg
C1	530	T 1 -	<u>۲۱.</u> .	C	Dec -	535	4	C1	D.c.	DI.	540	n	c.	U. 1	
	МΓХ	116	oru	Ser		se r	arg	uly	rro		Ala	rro	ser	val	
545 Cla	Tv	61 o	Cla	Ç	550	Dance	Arm.	Cv~	Т1,	555	41-	1	C1~	A	560
OTH	1 y I	uly	913	Ser 565	arg	110	нsр	LYS		110	HIB	ren	oin		
				ンいつ			•		570					575	

```
<;211>: 1728
 <:212>: DNA
 <;213>; Oryzias lapites
 <:220>:
 <:221>; CDS
 <:222>: (1)...(1728)
<:400>: 2
alg lac cel gaa gag age egg ggt tet gga ggg gtg get get gtg gae
Met Tyr Pro Glu Glu Ser Arg Gly Ser Gly Gly Val Ala Ala Val Asp
                  5
                                     10
ett tig gan ggg acg tac gac tat gee gee eec aac eet gee acg act
Leu Leu Glu Gly Thr Tyr Asp Tyr Ala Ala Pro Asn Pro Ala Thr Thr
             20
                                  25
occ ett tac age cag tee age acc gge tac tac tet get eec etg gaa
Pro Leu Tyr Ser Gla Ser Ser Thr Gly Tyr Tyr Ser Ala Pro Leu Glu
aca aac gga ccc ccc toa gaa ggc agt ctg cag toc ctg ggc agt ggg
                                                                    192
Thr Ash Gly Pro Pro Ser Glu Gly Ser Leu Gln Ser Leu Gly Ser Gly
ceg acg age cot etg gtg ttt gtg ecc tec age ecc aga etc agt ecc
                                                                    240
Pro Thr Ser Pro Leu Val Phe Val Pro Ser Ser Pro Arg Leu Ser Pro
 65
ttt atg cat coa eec age cac cac tat etg gaa acc act tee acg eec
                                                                    288
Phe Met His Pro Pro Ser His His Tyr Leu Glu Thr Thr Ser Thr Pro
                 85
                                     90
gtt tac aga tec age cac cag gga gee tec agg gag gac cag tge gge
                                                                    336
Val Tyr Arg Ser Ser His Gln Gly Ala Ser Arg Glu Asp Gln Cys Gly
            100
                                105
tee egg gag gac acg tge age etg ggg gag tta gge gee gga gee ggg
                                                                    384
Ser Arg Glu Asp Thr Cys Ser Leu Gly Glu Leu Gly Ala Gly Ala Gly
                            120
get ggg ggg tit gag atg gee aaa gae acg egt tie tge gee gig tge
                                                                    432
Ala Gly Gly Phe Glu Met Ala Lys Asp Thr Arg Phe Cys Ala Val Cys
                        135
age gae tae gee tet ggg tae eae tat ggg gtg tgg tet tgt gag gge
                                                                    480
Ser Asp Tyr Ala Ser Gly Tyr His Tyr Gly Val Trp Ser Cys Glu Gly
                                        155
                    150
tgc aag ged tto tto aag agg age ato cag ggt cac aat gad tat atg
Cys Lys Ala Phe Phe Lys Arg Ser Ile Gln Gly His Asn Asp Tyr Met
                165
                                    170
tge cea geg ace aat eag tge act att gae aga aat ega agg aag gge
                                                                    576
Cys Pro Ala Thr Asn Gln Cys Thr He Asp Arg Asn Arg Arg Lys Gly
            180
                                185
tgt cag get tgt egt ett agg aag tgt tac gaa gtg gga atg atg aaa
                                                                    624
Cys Gln Ala Cys Arg Leu Arg Lys Cys Tyr Glu Val Gly Met Met Lys
        195
                            200
gge ggt gtg ege aag gae ege att ege att tta egg egt gae aaa egg
                                                                    672
Gly Gly Val Arg Lys Asp Arg 11e Arg 11e Leu Arg Arg Asp Lys Arg
```

	210	С	•			215	;				220	1				
cg	g aca	a ggc	gtt	ggt	gat	gga	gac	aag	gtt	gta	aas	ggt	cag	gag	cat	720
Arg	g Thi	r Gly	y Val	Gly	/ Asp	Gly	Asp	Lys	. Val	Val	Lys	Gly	Gln	Glu	His	
225	5				230)				235	,		•		240	
aaa	ac;	g gts	g cat	tat	gat	gga	agg	aaa	cgc	age	agc	aca	gga	gga	gga	768
													*		Gly	
		-		245					250					255		
gga	ı gga	ı ggə	gga	gga	gga	aga	ctg	tet	gtg	acc	agc	ata	cct	cct	gag	816
			Gly													
			260		٠.			265					270			
cas	gts	cts	ctc	ctc	ctt	cag	ggc	goo	gag	ccc	cog	ata	ctc	tgc	tcg	864
			Leu													
		275					280					285		-,		
cgt	. cag		ttg	age	cga	ccg			ซลซ	etc	acc		atg	acc	cté	912
															Leu	714
,	290			2.01		295				, 441	300			1111	LCu ,	
cte			atg	gca	gae		(/a (/	et a	ote	cac		atc.	gee	taa	acc	960
			Met													900
305		JCI	110.0	nia.	310	Lys	U1u	Leu	vai	315	TICLU	1.10	AIG	11 1	320	
		cto	cca	aat		cta	/·26	eta	100		CaC.	u a t		ata		1000
			Pro													1008
Lys	LJS	Leu	110	325	rne	Leu	0111	Leu		Leu	1115	ASP	GIH		Leu	
a ta	ot a	do.d	240		l aa	a+ a	:		330	.+	212			335		1050
			agc													1056
Leu	Leu	GIU	Ser	ber	тр	Leu	GIU			met	пе	GIY			ПЪ	
	4		340					345		4.4			350			
			cac													1104
arg	Ser		His	Cys	Pro	GI y		Leu	He	Phe	Ala		Asp	Leu	He	
		355					360					365		_		
			aat													1152
Leu		Arg	Asn	Glu	Gly		Cys	Val	Glu	Gly		Thr	Glu	He	Phe	
	370					375					380					
			ctg													1200
	Met	Leu	Leu	Ala	Thr	Ala	Ser	Arg	Phe	Arg	Val	Leu	Lys	Leu	Lys	•
385					390					395					400	. :
			ttc													1248 .
'ro	GI u	Glu	Phe	Val	Cys	Leu	Lys	Ala	He	He	Leu	Leu	Asn	Ser	Gly	
				405					410					415		
			ttc													1296
Ma	Phe	Ser	Phe	Cys	Thr	Gly	Thr	Met	Glu	Pro	Leu	His	Asn	Ser	Ala	
			420					425					430			
gcg	gtt	cag	agc	atg	ctg	gac	acc	atc	aca	gac	gca	ctc	att.	cat	tac	1344
lla	Val	Gln	Ser	Met.	Leu	Asp	Thr	He	Thr	Asp	Ala	Leu	He	His	Tyr	
		435					440					445				
itc	agt	cag	t.cg	ggt.	tac	ttg	gee	cag	gag	cag	geg	aga	egg	cag	gcc	1392
			Ser													-
	450					455					460					
ag		ale	ctg	ct.g			cac	at.c	agg			agc	aac	aaa	ggc	1440
			Leu I													
65		,			470					475		•	-	-	480	
	gag	cac	ctc			at.c	aag	tec	aag		ааа	gtc	ect.	ct1.		1488
		🕶					0	٠.٠.				500	0		3.50	11.0

```
Met Glu His Leu Tyr Ser Met Lys Cys Lys Asn Lys Val Pro Leu Tyr
                485
                                     490
gae etc eta etg gag atg etc gat gec cac ege etg cac cac ece gte
Asp Leu Leu Clu Met Leu Asp Ala His Arg Leu His His Pro Val
                                505
aga gee een eag tee tig tee caa gie gae aga gae eet een tee acc
                                                                   1584
Arg Ala Pro Glin Ser Leu Ser Glin Val Asp Arg Asp Pro Pro Ser Thr
        515
                             520
                                                 525
age age gge ggg ggt gga ate get eee ggt tet ataltea gea tet ega
                                                                   1632
Ser Ser Gly Gly Gly He Ala Pro Gly Ser He Ser Ala Ser Arg
                        535
gge aga ate gag agt eeg age aga gge eec ttt get eee agt gte ett
                                                                   1680
Gly Arg He Glu Ser Pro Ser Arg Gly Pro Phe Ala Pro Ser Val Leu
                    550
                                        555
cag tat gga ggg tog ogt oct gac tgc acc ocg gcc ott caa gac tga
                                                                   1728
Gln Tyr Gly Gly Ser Arg Pro Asp Cys Thr Pro Ala Leu Gln Asp
                565
                                    570
<;210>: 3
<;211>: 620
<;212>; PRT
<;213>; Oryzias lapites
<:400>: 1
Met Ser Lys Arg Glin Ser Ser Val Glin He Arg Glin Leu Phe Gly Pro
Ala Leu Arg Ser Arg He Ser Pro Ala Ser Ser Glu Leu Glu Thr Leu
                                 25
Ser Pro Pro Arg Leu Ser Pro Arg Asp Pro Leu Gly Asp Met Tyr Pro
Glu Glu Ser Arg Gly Ser Gly Gly Val Ala Ala Val Asp Leu Leu Glu
                         55
Gly Thr Tyr Asp Tyr Ala Ala Pro Asn Pro Ala Thr Thr Pro Leu Tyr
Ser GIn Ser Ser Thr Gly Tyr Tyr Ser Ala Pro Leu Glu Thr Asn Gly
                 85
                                     90 -
Pro Pro Ser Glu Gly Ser Leu Gln Ser Leu Gly Ser Gly Pro Thr Ser
                                105
Pro Leu Val Phe Val Pro Ser Ser Pro Arg Leu Ser Pro Phe Met His
                            120
                                             125
Pro Pro Ser His His Tyr Leu Glu Thr Thr Ser Thr Pro Val Tyr Arg
                                            140
                        135
Ser Ser His Glu Gly Ala Ser Arg Glu Asp Glu Cys Gly Ser Arg Glu
                    150
                                        155
Asp Thr Cys Ser Leu Gly Glu Leu Gly Ala Gly Ala Gly Ala Gly Gly
                                    170
Phe Glu Met Ala Lys Asp Thr Arg Phe Cys Ala Val Cys Ser Asp Tyr
                                185
Ala Ser Gly Tyr His Tyr Gly Val Trp Ser Cys Glu Gly Cys Lys Ala
                            200
                                                205
```

Phe	Phe 210		Arg	Ser	He	G1n 215		His	Asn	Asp	Tyr 220	Met	Cys	Pro	Ala
Thr 225		Gln	Cys	Thr	11e 230	Asp	Arg	Asn	Arg	Arg 235	Lys	Gly	Cys	Gln	Ala 240
		Leu	Arg			Tyr	Glu	Val			Met	Lys	Gly		
Arg	, Lys	Asp	Arg	245 11e	Arg	lle	Leu	Arg	250 Arg	Asp	Lys	Arg	Arg	255 Thr	Gly
0.1	CI.	• • • • • • • • • • • • • • • • • • •	260	A	l	V - 1		265	C1	C1 n	C1	III.a	270		Val
		275	Gly		٠.		280		•	:		285			
llis	Tyr 290		Gly	Arg	Lys	Ang 295	Ser	Ser	Thr	Gly	G1y 300	Gly	Gly	Glý	Gly
Gly			Arg	Leu	Ser		Thr	Ser	He	Pro		Glu	Gln	Val	Leu
305		·		•	310	:			,	315			,		320
Leu	Leu	Leu	GIn	Gly	Ala	Glu	Pro	Pro	He	Leu	Cys	Ser	Arg	Gln	Lys
		•		325					330					335	
Leu	Ser	Arg	Pro 340		Thr		Val	Thr 345	Met	Met.	Thr	Leu	Leu 350	Thr	Ser
Met	Ala	Asp 355	Lys	Glų	Leu	Val	His 360		He		Trp	Ala 365	Lys	Lys	Leu
Pro	Glv		Leu	Gľn	Leu	Ser					Val		Leu	Leu	Glu
	370		7.7			375					380				
Ser	Ser	Trp	Leu	Glu	Val	Leu	Met	He	Gly	Leu	He	Trp	Arg	Ser	lle
385					390					395					400
His	Cys	Pro	Gly	Lys	Leu	He	Phe	Ala	Gln	Asp	Leu	He	Leu	Asp	Arg
		•		405					410					415	
Asn	Glu	Gly	Asp					_				Phe		Met	Leu
Lau	41.0	Th.	420 Ala		٨٠٠٠			425		l ve		l ve	430	Cl.,	Clu
Leu	Ala	435	MId	Jei		The	440	vai	Leu	Lys	Leu	445	. 10	oru	aru
Phe	Val		Leu	Lys	Ala,	He		Leu	Leu	Asn	Ser		Ala	Phe	Ser
	450			٠.		455		•			460				
Phe	Cys	Thr	Gly	Thr	Met	Glu	Pro	Leu	llis	Asn	Ser	Ala	Ala	Val	Gln
					470										
Ser	Met	Leu	Asp	Thr _.	He	Thr	Asp	Ala	Leu 490	He	His	Tyr	He	Ser 495	Gln
Ser	Gly	Tyr	Leu	Ala	Gln	Glu	GIn		Arg	Arg	Gln			Pro	Leu
			500					505	C		•		510	C1	112 -
Leu	Leu	Leu 515	Ser	HIS	116	Arg	нтs 520	Met	ser	ASN		525	met	GIU	HIS
Leu	Tyr 530	Ser	Met	Lys	Cys	Lys 535	Asn	Lys	Val	Pro	Leu 540	Туг	Asp	Leu	Leu
Leu	Glu	Met	Leu	Asp	Ala	His	Arg	Leu	His	His	Pro	Val	Årg	Ala	Pro
545					550					555	٠.				560
Gln	Ser	Leu	Ser	G1 n 565	Val	Asp	Arg	Asp	Pro 570	Pro	Ser	Thr	Ser	Ser 575	Gly
Gly	Gly	Gly	He		Pro	Gly	Ser	He		Ala	Ser	Arg	Gly		Пe
-	-		580					585					590		
Glu	Ser	Pro	Ser	Arg	Gly	Pro	Phe	Ala	Pro	Ser	Val	Leu	Gln	Tyr	Gly
		595					600					605			

Gly Ser Arg Pro Asp Cys Thr Pro Ala Leu Gln Asp	
610 615	
<;210≥; 2	
<;211>; 1863	
<:21⊅; DNA	
<;213>; Oryzias lapites	
<:220>;	
<;221>; CDS	
<;22⊅; (1)(1863)	
<;400>; 2	
ats agt aas aga cas age tog sts cas atc ags cas cts ttc sga cca	48
Met Ser Lys Arg Gln Ser Ser Val Gln He Arg Gln Leu Phe Gly Pro	
1 5 10 15	
gea etc aga tec agg atc age cea gee tec tea gag etg gag acc etc	96
Ala Leu Arg Ser Arg Ile Ser Pro Ala Ser Ser Glu Leu Glu Thr Leu	
20 25 30	
too coa cet ege etc tog eee egt gae eee etc ggt gae atg tac eet	144
Ser Pro Pro Arg Leu Ser Pro Arg Asp Pro Leu Gly Asp Met Tyr Pro	
35 40 45	
gaa gag age egg ggt tet gga ggg gtg get get gtg gae ett ttg gaa	192
Glu Glu Ser Arg Gly Ser Gly Gly Val Ala Ala Val Asp Leu Leu Glu	
50 55 60	
ggg acg tac gac tat gcc gcc ccc aac cct gcc acg act ccc ctt tac	240
Gly Thr Tyr Asp Tyr Ala Ala Pro Asn Pro Ala Thr Thr Pro Leu Tyr	
65 70 75 80	
	288
Ser Gln Ser Ser Thr Gly Tyr Tyr Ser Ala Pro Leu Glu Thr Asn Gly	
85 90 95	
	336
Pro Pro Ser Glu Gly Ser Leu Gln Ser Leu Gly Ser Gly Pro Thr Ser	
100 105 110	
cet etg sts ttt sts eec tee age eec aga etc agt eec ttt atg eat	384
Pro Leu Val Phe Val Pro Ser Ser Pro Arg Leu Ser Pro Phe Met His	
115 120 125	
	432
Pro Pro Ser His His Tyr Leu Glu Thr Thr Ser Thr Pro Val Tyr Arg	
130 135 140	
	180
Ser Ser Ilis Gln Gly Ala Ser Arg Glu Asp Gln Cys Gly Ser Arg Glu	
145 150 155 160	
	528
Asp Thr Cys Ser Leu Gly Glu Leu Gly Ala Gly Ala Gly Ala Gly Gly	- -
165 170 175	
	576
	<i>,</i> 10
Phe Glu Met Ala Lys Asp Thr Arg Phe Cys Ala Val Cys Ser Asp Tyr 180 185 190	
	624
ged tot ggg tad dad tat ggg gtg tgg tot tgt gag ggd tgd aag gdd 1	J44 -

	Λla	Ser	Gly 195	Tyr	His	Tyr	Gly	Val 200	Trp	Ser	Cys	G1 u	G1y 205	Cys	l.ys	Ala	
	tte	ite	aag	agg	agc	atc	cag	ggt	cac	aat	gac	tat.	atg	tgc	cca	gcg	672
						He											•
•		210		3			215				•	220					
				4		- 44			a a #	000	5 aa		aa0	1.4	00 G	ant	720
						att		-									1.20
		Asn	GIN	Uys	m	He	ASP	Arg	ASII	A) g		Lys	иту	Lys	OH		
	225					230					235				. :	240	
						tgt.											768
	Cys	Arg	Leu	Arg	Lys	Cys	Tyr	Gļu	Val	Gly	Met	Met	Lys	Gly	Gly	Val	
-					245					250					255	•	
	cgc	aag	gac	cgc	at t	cgc	att	tta	egg	cgt	gac	aaa	cgg	cgg	aca	ggc	816
	Arg	Lys	$\overline{\mathrm{Asp}}$	Arg	Пе	Arg	He	Leu	Arg	Arg	Asp	Lys	Arg	Arg	Thr	Gly	
				260					265					270			
	gtt	ggt	gat	gga	gac	aag	gtt	gta	aag	ggt	cag	gag	cat	aaa	acg	gtg	864
						Lys											
			275					280		·			285	Ì			
	cat	tat	gat	gga	agg	aaa	cgc	agc	agc	aca	gga	gga	gga	gga	gga	gga	912
	His	Tyr	Asp	${\rm Gl}{\bf y}$	Arg	Lys	Arg	Ser	Ser	Thr	Gly	Gly	Gly	Gly	Gly	Gly	
		290					295					300					
	gga	gga	gga	aga	ctg	tct	gtg	acc	agc	ata	cct	cct	gag	cag	gtg	ctg	960
	Gly	Gly	Gly	Arg	Leu	Ser	Val	Thr	Ser	He	Pro	Pro	Glu	Gln	Val	Leu	
	305	-		1. 1.		310					315					320	
	ctc	cte	ctt	cag	ggc	gcc	gag	ссс	cog	ata	ctc	tgc	teg	cgt	cag	aag	1008
						Ala											÷
					325					330		•		-	335	Ϊ,	
	He	300	cda	ees.		acc	03 g	etc	acc		atg	acc	cte	ctc		age	1056
						Thr											
	L.v.u			340		1 111	1,114	, 41	345	18.0	710.0		1.2.0	350	••••	501	
	n t a	uter n		-		ctg	ate			ato	unc	+ aa	1100		220	cte	1104
																	1104
	met.	Ara	355	Lys	GIU	l.eu	vai	360	met	He	Ala	пр	365	1.35	Lys	Leu	
	cca	ggt	ttt	ctg	cag	ctg	tcc	ctg	cac	gat	cag	gtg	ctg	ctg	ctg	gag	1152
						Leu											
	-	370					375			_		380	•				
	age		t erer	ctø	ខ្លួក	gtg		ate	atc	୧୧୯	ctc		t.ee	agg	tee	atc	1200
						Val											
			ii b	1.4.4	· · · · · ·	390	12.0	nee	110.	41,3	395					400	
	385	11					.4	+++	a	005		o+	ata.	ot a	420		1248
						ctc											1240
	nis	Lys	PTO	uly		Leu	пе	rne	Ala		asp	Leu	He	Leu		Arg	
					105					410					415		
						gtg											1296
	Asn	Glu	Gly	Asp	Cys	Val	Glu,	Gly	Met	Thr	Glu	Пе	Phe	Asp	Met.	L.eu	
				420					425			٠.		430			
	ctg	gcc	act	get	tec	ogo	ttc	cgt	gtg	ctc	aaa	ctc	aaa	cct	gag	gaa	1344
	l.eu	Ala	Thr	Ala	Ser	Arg	Phe	Arg	Val	Leu	Lys	Leu	Lys	Pro	$G1\mathbf{u}$	Glu	
			435					440					445				
	ttc	gtc		ctc	aaa	gct	att	att	tta	ctc	aac	tee	ggt	get	ttt	tct	1392
						Ala											
		450	- • •				455					460	•				
		150					200										

tte	tgc	acc	ggc	acc	atg	gag	сса	ctt	cac	aac	age	gcg	gcg	gt.t.	cag	1/440
Phe	Cys	Thr	Gly	Thr	Met	Glu	Pro	Leu	llis	Asn	Ser	Ala	Ala	Va I	G1n >	
465					470					475	•. •				480	
age	atg	ctg	gac	acc	ate	aca	gac	gca	ctc	att	cat.	tac	atc	agt.	cag	1488
Ser	Met	Leu	Asp	Thr	He	Thr	Asp	Ala	Leu	He	llis	Tyr	He	Ser	Gln	•
	•			485					490					495		
tcg	ggt	tac	t.t.g	gcc	cag	gag	cag	gcg	aga	cgg	cag	gee	cag	ccg	ctc	1536
Ser	Gly	Туг	Leu	Ala	Gln	Glu	Gln	Ala	Arg	Arg	Gln	Ala	Gln	Pro	Leu	
			500					505				٠:	510			
ctg	ctg	ctc	tcc	cac	atc	agg	cac	atg	agc	aac	aaa	ggc	atg	gag	cac	1584
Leu	Leu	Leu	Ser	llis	He	Arg	His	Met	Ser	Asn	Lys	Gly	Met	Glu	His	
	•	515		•			520					525				
ctc	tac	age	atg	aag	tgc	aag	aac	aaa	gte	cct	ctl	tat	gac	ctc.	cta	1632
Léu	Tyr	Ser	Met.	l.ys	Cys	Lys	Asn	Lys	Val	Pro	Leu	Tyr	Asp	Leu	Leu	
	530		•			535					540					•
ctg	gag	atg	ctc	gat	gcc	cac	cgc	ctg	cac	cac	ccc	gtc	aga	gcc	ccc	1680
Leu	Glu	Met	Leu	Asp	Ala	His.	Arg	Leu	His	His	Pro	Val	Arg	Ala	Pro	
545	• 7				550					555					560	
cag	tec	ttg	tcc	caa	gte	gac	aga	gac	cct	ccc	tee	acc	agc	agc	gge	1728
ĠIn	Ser	Leu	Ser	Gln	Val	Asp	Arg	Asp	Pro	Pro	Ser	Thr	Ser	Ser	Gly	
				565					570					575		
ggg	ggt	gga	atc	gct	ccc	ggt	tct	ata	t.ca	gca	$t \alpha : t$	cga	ggc	aga	atc	1776
Gly	Gly	Gly	Пe	Ala	Pro	Gly	Ser	Цe	Ser	Ala	Ser	Arg	Gly	Arg	He	
			580					585					590			
gag	agt	ccg	agc	aga	ggc	ccc	ttt	get	ccc	agt	gtc	ctt	cag	tat	gga	1824
Glu	Ser	Pro	Ser	Arg	Gly	Pro	Phe	Ala	Pro	Ser	Va 1	Leu	Gln	Tyr	Glý	
		595					600					605				
ggg	teg	cgt	cct	gac	tge	acc	ccg	gcc	ctt	caa	gac	t ga				1863
Gly	Ser	Arg	Pro	Asp	Cys	Thr	Pro	Ala	Leu	Gln	Asp					
	610					615					620					

【図面の簡単な説明】

【図1】本発明のエストロジェンレセプター遺伝子を発現させるための発現ベクターの構築過程を示す図である。

【図2】本発明のエストロジェンレセプター遺伝子を発現させるための発現ベクターRSV-ERの構造を示す図である。pRC/RSVは構築に用いたベクターである。ERは本発明のエストロジェン遺伝子を、RSV LTRはRSVプロモーターを、pSV40はSV40プロモーターを、Neomycinはネオマイシン耐性遺伝子をそれぞれ示す。

【図3】エストロジェン応答配列とレポーター遺伝子と を含むレポータープラスミドERE-tK-pGL(図中ではEREpGLと表示)の構築過程を示す図である。

【図4】エストロジェン応答配列とレボーター遺伝子とを含むレポータープラスミドERE-tK-pGLの構造を示す図である。4XEREはエストロジェン応答配列が4個連結された配列を意味し、tK promoterはtKプロモーターを、Ampでアンピシリン耐性遺伝子を示す。

【図5】レポーターアッセイ用の細胞を調製するための トランスフェクションの条件を検討した結果を示す図で ある。上段の図は、トランスフェクション試薬にリポフェクチンを使用した場合、下段の図はリポフェクトアミンを使用した場合の結果を示す。

【図6】レポーターアッセイ用の細胞を調製するためのトランスフェクションにおいて、添加するトランスフェクション試薬の量を検討した結果を示す図である。上段の図は、リボフェクチンを使用した場合、下段の図はリポフェクトアミンを使用した場合の結果を示す。

【図7】レポーターアッセイ用の細胞を調製するためのトランスフェクションにおいて、添加するプラスミドの量を検討した結果を示す図である。上段の図は、リポフェクチンを使用した場合、下段の図はリポフェクトアミンを使用した場合の結果を示す。

【図8】レボーターアッセイにおける β -エストラジオールのエストロジェンレセプター活性化能を測定した結果を示す図である。1ウェルあたり0.25 μ gのレセプター発現ベクターRSV-ER、0.15 μ gのレボータープラスミドERE5-tK-pGL、および、0.1 μ gのコントロールレボータープラスミドpRL-tKを導入した細胞を試験に用いた。上段の図は、トランスフェクション試薬にリポフェクチ

ンを使用した場合、下段の図はリボフェクトアミンを使 用した場合の結果を示す。

【図9】レボーターアッセイにおける β ーエストラジオールのエストロジェンレセプター活性化能を測定した結果を示す図である。1ウェルあたり0.25 μ gのレセプター発現ベクターRSV-ER、0.15 μ gのレポータープラスミドERE5-tK-pGL、および、0.1 μ gのコントロールレポータープラスミドpRL-tKをリポフェクタミンを用いて導入した細胞を試験に用いた。

【図10】レポーターアッセイにおけるビスフェノール Aのエストロジェンレセプター活性化能を測定した結果 を示す図である。1ウェルあたり0.25 μ gのレセプター 発現ベクターRSV-ER、0.15 μ gのレポータープラスミドE RE5-tK-pGL、および、0.1 μ gのコントロールレポーター プラスミドpRL-tKをリポフェクタミンを用いて導入した 細胞を試験に用いた。E2は、終濃度500pMの β -エストラジオールが添加された系を示す。

【図11】レポーターアッセイにおけるp-ノニルフェノールのエストロジェンレセプター活性化能を測定した結果を示す図である。1ウェルあたり0.25 μ gのレセプター発現ベクターRSV-ER、0.15 μ gのレポータープラスミドERE5-tK-pGL、および、0.1 μ gのコントロールレポータープラスミドpRL-tKをリポフェクタミンを用いて導入した細胞を試験に用いた。E2は、終濃度500pMの β -エストラジオールが添加された系を示す。

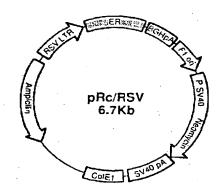
【図12】レポーターアッセイにおける酢酸トリブチルすずのエストロジェンレセプター活性化能を測定した結果を示す図である。1ウェルあたり 0.25μ gのレセプター発現ベクターRSV-ER、 0.15μ gのレポータープラスミドERE5-tk-pGL、および、 0.1μ gのコントロールレポー

タープラスミドpRL-tKをリボフェクタミンを用いて導入 した細胞を試験に用いた。E2は、終濃度500pMσ)β-エストラジオールが添加された系を示す。

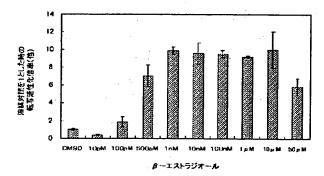
【図13】レボーターアッセイにおけるフタル酸ジー2 ーエチルへキシルのエストロジェンレセプター活性化能 を測定した結果を示す図である。1ウェルあたり0.25 μ gのレセプター発現ベクターRSV-ER、0.15 μ gのレボータ ープラスミドERE5-tK-pGL、および、0.1 μ gのコントロ ールレボータープラスミドpRL-tKをリポフェクタミンを 用いて導入した細胞を試験に用いた。E2は、終濃度50 0pMの β -エストラジオールが添加された系を示す。

【図14】レポーターアッセイにおける各種化合物のエ ストロジェンレセプター活性化能を測定した結果を示す 図である。1ウェルあたり0.25μgのレセプター発現べ クターRSV-ER、0.15μgのレポータープラスミドERE5-tK -pGL、および、0.1μgのコントロールレポータープラス ミドpRL-tKをリポフェクタミンを用いて導入した細胞を 試験に用いた。E 2は終濃度500pMのβ-エストラジオ ールが添加された系を、Tisは終濃度50μMの酢酸トリブ チルすずが添加された系を、BisAは終濃度50μMのビス フェノールAが添加された系を、Di-2は終濃度50μMの フタル酸ジー2ーエチルヘキシルが添加された系を、No ny1は終濃度50μMのノニルフェノールが添加された系を 示す。顕微鏡観察により判定された死亡細胞率は、酢酸 トリブチルすずが添加された系において90%以上、ビ スフェノールAが添加された系において0%、フタル酸 ジー2ーエチルヘキシルが添加された系において90% 以上、ノニルフェノールが添加された系において40% 前後であった。

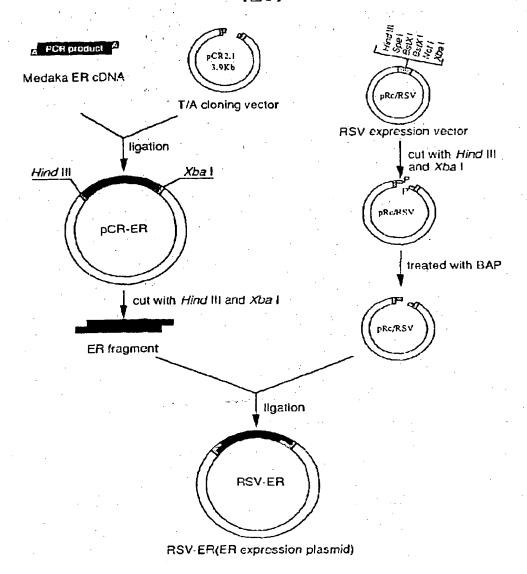
【図2】

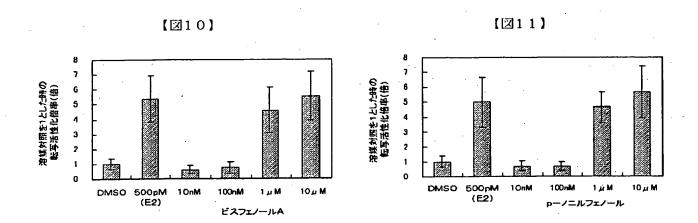


【図9】

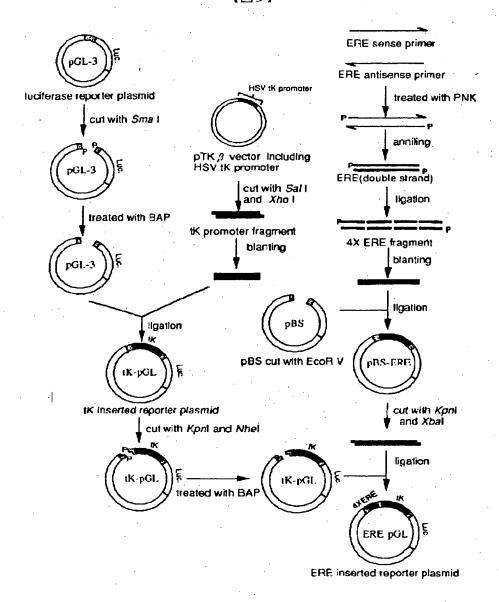


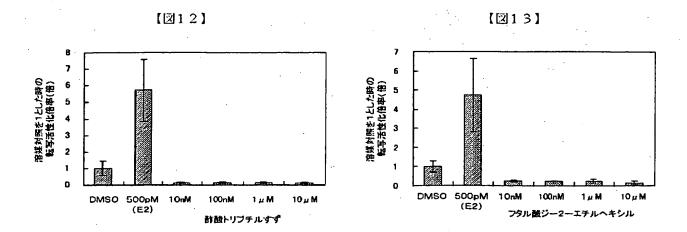
【図1】



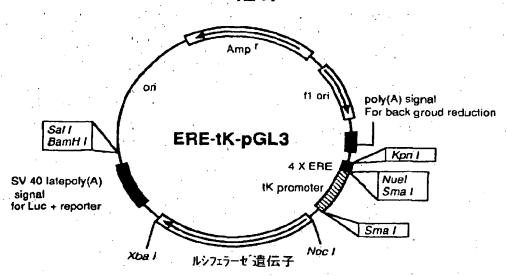


【図3】

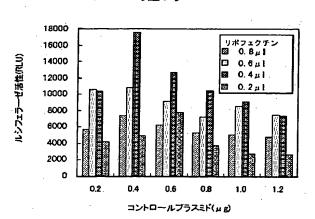












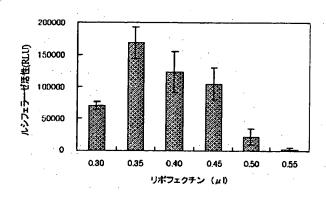
リポフェクタミン 図 0.8μ1 1400000 □ 0.6 µ 1 0.4 µ] 1200000 0. 2 μ 1 1000000 800000 600000 400000 200000

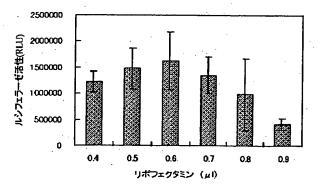
コントロールブラスミド(μg)

1.0

1.2

【図6】

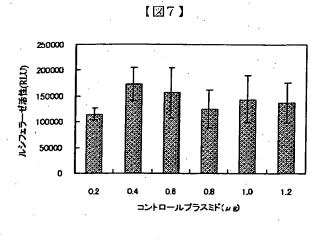


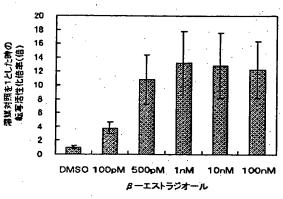


1800000

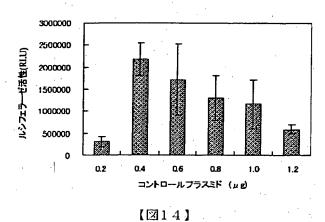
1600000

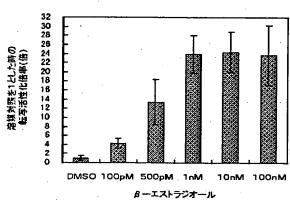
0.2

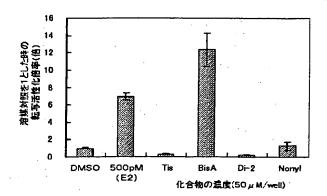




[図8]







フロントページの続き

 識別記号 ZNA

F ターム(参考) 4B024 AA11 AA17 AA20 BAG3 DA02 EA04 FA02 FA10 GA13 HA03 HA11

> 4B063 QA01 QQ22 QQ61 QQ75 QQ91 QQ94 QR33 QR60 QR77 QR80 QS36 QX02

> 4B064 AG20 CA10 CA19 CC24 DA13 DA16

4B065 AA90Y AA93X AB01 AC14 BA05 CA46

4H045 AA10 AA20 AA30 BA10 CA52 DA50 EA50 FA72 FA74 HA06

P.D. 05 -01 - 2001

, Ip .AAA92175 standard; DNA; 1863 BP.

AAA92175;

XP-002181483

XX

XX

AC

DT XX

DE XX

KW XX

OS XX

PN XX

PD XX

PF XX

PR XX

PA XX

DR

DR XX

PT XX

PS XX

CC

SQ

05-JAN-2001 (first entry)---

Oryzias lapites oestrogen receptor encoding DNA-SEQ-ID NO:4. -

Oryzias lapites; oestrogen receptor, ds.

Orygias lapites.

JP2000201688-A.

25-JUL-2000.

06-APR-1999; 99JP-0098787.

10-NOV-1998; 98JP-0319465.

(SUMO) SUMITOMO CHEM CO LTD.

WPI; 2000-567950/53. P-PSDB; AAB20898.

An estrogen receptor gene and its application \

Claim 4; Page 15-17; 23pp; Japanese.

The present sequence encodes an oestrogen receptor derived from Oryzias lapites. Also described are: (1) a vector comprising the oestrogen receptor gene; (2) a transformant prepared by introducing the oestrogen receptor gene or vector from (1) into a host cell; (3) a method for the preparation of an oestrogen receptor comprising culturing the transformant from (2) to produce the oestrogen receptor; and (4) a method for the evaluation of oestrogen receptor-activating ability of a chemical substance in which the chemical substance is reacted with a transformant prepared by introducing a reporter gene connected downstream of a transcription controlling region containing an oestrogen response sequence and the above oestrogen receptor gene to an oestrogen-nonendogenous host cell. The transformant can be used for the evaluation of oestrogen receptor-activating ability of a chemical substance.

Sequence 1863 BP; 406 A; 565 C; 531 G; 361 T; 0 other; atgagtaaga gacagagete ggtgcagate aggcagetgt teggaceage acteagatee 60 aggatcagec cagectecte agagetggag accetetece cacetegeet etegeceegt 120 __gacccctcg_gtgacatgta ccctgaagag agccggggtt ctggaggggt ggctgctgtg 180 240 gaccttttgg aagggacgta cgactatgcc gcccccaacc ctgccacgac tcccctttac agccagtcca gcaccggcta ctactctgct cccctggaaa caaacggacc cccctcagaa 300 ggcagtctgc agtccctggg cagtgggccg acgagccctc tggtgtttgt gccctccagc 360 cccagactca gtccctttat gcatccaccc agccaccact atctggaaac cacttccacg 420 480 cccgtttaca gatccagcca ccagggagcc tccagggagg accagtgcgg ctcccgggag gacacgtgca gcctggggga gttaggcgcc ggagccgggg ctggggggtt tgagatggcc 540 600 aaagacacgc gtttctgcgc cgtgtgcagc gactacgcct ctgggtacca ctatggggtg 660 tggtcttgtg agggctgcaa ggccttcttc aagaggagca tccagggtca caatgactat 720 atgtgcccag cgaccaatca gtgcactatt gacagaaatc gaaggaaggg ctgtcaggct tgtcgtctta ggaagtgtta cgaaqtggga atgatgaaag gcggtgtgcg caaggaccgc 780 attcgcattt tacggcgtga caaacggcgg acaggcgttg gtgatggaga caaggttgta 840 aagggtcagg agcataaaac ggtgcattat gatggaagga aacgcagcag cacaggagga 900 ggaggaggag gaggaggagg aagactgtct gtgaccagca tacctcctga gcaggtgctg 960 ctcctccttc agggcgccga gcccccgata ctctgctcgc gtcagaagtt gagccgaccg 1020 tacaccgagg tcaccatgat gaccctgctc accagcatgg cagacaagga gctggtccac 1080 atgatogoot gggccaagaa gotoccaggt tttotgcago tgtocctgca cgatcaggtg 1140

ctgctgctgg	agagctcgtg	gctggaggtg	ctcatgatcg	gcctcatttg	gaggtccatc	. 1200
		ctttgcacaa				1260
		gatcttcgac				1320
		ggaattcgtc				1380
		cggcaccatg				1440
agcatgctgg	acaccatcac	agacgcactc	attcattaca	tcagtcagtc	gggttacttg	1500
gcccaggagC	aggcgagacg	gcaggcccag	ccgctcctgc	tgctctccca	catcaggcac	1560
		gcacctctac				1620
tatgacctcc	tactggagat	gctcgatgcc	caccgcctgc	accaccccgt	cagagccccc	1680
cagtccttgt	cccaagtcga	cagagaccct	ccctccacca	gcagcggcgg	gggtggaatc	1740
gctcccggtt	ctatatcagc	atctcgaggc	ag aatega ga	gtccgagcag	aggccccttt	1800
gctcccagtg	tccttcagta	tggagggtcg	cgtcctgact	gcaccccggc	ccttcaagac	1860
tga	•					1863
-						

25-10-2001 15:

```
initn: 9261 init1: 9261 opt: 9261 Z-score: 8922.0 bits: 1664.7 E():
99.678% identity (99.678% ungapped) in 1863 nt overlap (76-1938:1-1863)
              60
                      70
                              80 .
                                      90
EP0111 CGTGTTGCGCAGCACATCTGAGGATGATTCATGAGTAAGAGACAGAGCTCGGTGCAGATC
                           ..........
                            ATGAGTAAGAGACAGAGCTCGGTGCAGATC
GSN: AA
                                  10
                             140
       110
              120
                     130
EP0111 AGGCAGCTGTTCGGACCAGCACTCAGATCCAGGATCAGCCCAGCCTCCTCAGAGCTGGAG
     50
                          60
           40.
                             200
                     190
              180
      170
EP0111 ACCCTCTCCCCACCTCGCCTCTCGCCCCGTGACCCCCTCGGTGACATGTACCCTGAAGAG
     GSN: AA ACCCTCTCCCCACCTCGCCTCTCGCCCCGTGACCCCCTCGGTGACATGTACCCTGAAGAG
                                 130
                                                150
          100
                 110
                         120
                     250
              240
                             260
      230
EP0111 AGCCGGGGTTCTGGAGGGTGGCTGCTGTGGACTTTTTGGAAGGACGTACGACTATGCC
     GSN: AA AGCCGGGGTTCTGGAGGGGTGGCTGCTGTGGACCTTTTGGAAGGGACGTACGACTATGCC
                                190
                                         200
          160
                  170
                         180
                     310
                             320 `
                                     330 '
              300
EP0111 GCCCCCAACCCTGCCACGACTCCCCTTTACAGCCAGTCCAGCAGCGGCTACTACTCTGCT
     GSN: AA GCCCCCAACCCTGCCACGACTCCCCTTTACAGCCAGTCCAGCACCGGCTACTACTCTGCT
                                 250
                                        260
                                                270
                  230
                         240
          220
                                     390
                     370
                             380
              360
EP0111 CCCCTGGAAACAAACGGACCCCCCTCAGAAGGCAGTCTGCAGTCCCTGGGCAGTGGGCCG
     GSN: AA CCCCTGGAAACAAACGGACCCCCCTCAGAAGGCAGTCTGCAGTCCCTGGGCAGTGGGCCG
                         300 ...
                                 310
                                         320
                  290
          280
                      430
                             440
                                     450
              420
       410
EP0111 ACGAGCCCTCTGGTGTTTGTGCCCTCCAGCCCCAGACTCAGTCCCTTTATGCATCCACCC
     GSN: AA ACGAGCCCTCTGGTGTTTGTGCCCTCCAGCCCCAGACTCAGTCCCTTTATGCATCCACCC
                         360
                                 370
                  350
          340
       470
              480
                     490
                             -500 ·
EP0111 AGCCACCACTATCTGGAAACCACTTCCACGCCCGTTTACAGATCCAGCCACCAGGGAGCC
     GSN: AA AGCCACCACTATCTGGAAACCACTTCCACGCCCGTTTACAGATCCAGCCACCAGGGAGCC
                         420
                                 430
                  410
          400
                                     570
                      550
                             560
       530
EP0111 TCCAGGGAGGACCAGTGCGGCTCCCGGGAGGACACGTGCAGCCTGGGGGAGTTAGGCGCC
     GSN: AA TCCAGGGAGGACCAGTGCGGCTCCCGGGAGGACACGTGCAGCCTGGGGGAGTTAGGCGCC
                         480 -
                                 490
                                        500
                  470
          460
                                     63.Q
              600
                      610 . . .
                             620
EP0111 GGAGCCGGGGCTGGGGGGTTTGAGATGGCCAAAGACACGCGTTTCTGCGCCGTGTGCAGC
     GSN: AA GGAGCCGGGGCTGGGGGGTTTGAGATGGCCAAAGACACGCGTTTCTGCGCCGTGTGCAGC
                                         560
                                                570
          520
                  530
                         540
                                 550
                                     69.0
                                            700
       650
              660
                      670
                             680
```

>>GSN:AAA92175 Oryzias lapites oestrogen recept (1863 nt)

EPUIII		CGCCTC										
		:::::										
GSN: AA	GACTA		TGGGT		TATG		rggtc		AGGGC		GGCCT	
		580		590		60 0		610		620		630
•											•	
	710				730		740		750		760	•
EP0111												
		:::::										
GSN: AA	AAGAG	GAGCAT	CCAGG	GTCAC	AATG	CTATA	TGTG	CCCAG	CGACC	AATCA(STGCAC	TTATT
		640		650		660		670		680		690
٠.	770		780		790		800		810		820	
EP0111	GACAG.	AAATCĠ	GAGGA	AGAGC	TGCC	GGCTI	'GTCG'	TCTTA	GAAG	TGTTAC	CGAAGI	GGGA
. •		:::::										
GSN: AA												
		700		710		720		730		740		750
												, 50
•	. 83.0		840		850		860		870		880	
EP0111		-		тесес			-	<u>ጉ አ</u> ጥጥጥ ሰ		ാഗ്രസ്ഥ് മ		
DI VIII												
CCNAAA		:::::										
GSN: AA	AIGAI		CGGTG		AAGGA		TTCG				AAACG	
	• •	760		770	•	7 80		790		800		810
		•										
	89 0		900		910		920		930		940	
EP0111	ACAGG	CGTTGG	TGATG	GAGAC	aaggi	TGTAA	AGGG:	rcagg <i>i</i>	GCAT	AAAACG	GTGCA	TATT
		:::::										
GSN: AA	_ACAGG(CGTTGG	rgatg(GAGAC	AAGGI	TGTAA	AGGG:	CAGG	GCAT	AAAACG	GTGCA	TTAT
	•	820		830		840		85 0	•	860		870
											•	
	950	•	960		970		980	-	9 90		1000	
EP0111	GATGG	AAGGAAA	ACGCA	GCAGC	ACAGG	AGGAG	GAGG	GGAGG	AGGA	GAGGA	AGACT	GTCT
		: : : : : :										
GSN: AA												
		880		890		900		910		920		930
		¥. •						, , ,		7 - 7		230
	1010	1	1020	1	1030		1040		1050		1060	
EP0111												משמט
GSN: AA												
GDN.AA	GIGACC	940	100100	950	~4001	960 -			الالالالالالا		الدلدلال	
		940		950		960		970	, ,	7980		990
-	1070			-			1100					
DD0111	1070		1080		1090.				1110.		1120	
EP0111	CTCTGC											
		::::::										
GSN:AA	CTCTGC											
		1000	3	1010		1020	-	1030		1040		1050
												.*
	1130		140		150	_	1160		1170		1180	
EP0111												
		:::::										
GSN: AA	ACCAGO	ATGGCA	GACAA	GGAGC	TGGT	CCACA	rgate	GCCTG	GGCC	AGAAG	CTCCC	AGGT.
•		1060	-, 1	1070		1080		1090		1100	*	1110
	1190	1	200	1	210		1220		1230		1240	4
EP0111	TTTCTG	CAGCTG	TCCCI	GCACG	ATCA	GGTGC	rgcre	CTGGA	GAGCT			CCTC
		:::::										
GSN: AA												
GDIV.AA	111010	1120					r GC I G		GAGCI			
		1120	1	130		1140		1150		1160		1170
		_		_								
	1250	-	260		270	-	1280		1290		1300	
EP0111												
		:::::	:::::	:::::								
CC31 - 3.3												
GSN: AA	CTCATG	ATCGGC			GGTC		CTGT		GAAGO			

	.1310	_1320									
EP0111	GACCTCATC										
•	::::::::										
GSN: AA	GACCTCATC	CTGGAC	AGGAA!	rgagg	BAGACT	r cc ct	3GAAG6	CATG	CGGAG	ATCTT	CGAC
	124	0	1250		1260		1270		1280	-	1290
										•	
	1370	1380		1390		1400		1410		1420	
EP0111	ATGCTGCTG	GCCACT	CTTC	CCGCT	rccgr	TGCT	CAAACI	CAAAC	CTGAG	GAATT	CGTC
2.02	::::::::										
CCM. AA	ATGCTGCTG										
GSMLAA	130						1330				1350
	1,50		1310		1320		-550		1310		1000
	1430	1440		~1450		13 CA		9.770		1/00	
EPUILI	TGCCTCAAA										
	:::::::::										
GSN: AA	TGCCTCAAA	GCTATT	ATTTT7	ACT CAZ	ACTCC	GIGC.		TITC	CACC	GGCAC	
	1360	0	1370.		1380		1390		1400		1410
4.5	1490										
EP0111	GAGCCACTT						-				
	::::::::										
GSN:AA	GAGCCACTT	CACAAC	AGCGC	3GCGG:	PTCAGA	GCAT(CTG G2	CACCE	TCACA	GACG	ACTC
	1420	0	1430		1440		1450		1460		1470
	1550	1560		1570		1580		1590		160Q	
EP0111	ATTCATTAC	ATCAGT	CAGTC	GGTT	ACTTG	CCCA	GAGCA	GGCGA	GACGG	CAGGO	CCAG
	::::::::	::::::	:::::	:::::	:::::	:::::	:::::	:::::	:::::	:::::	:::::
GSN: AA	ATTCATTAC										
	1486	9 ·	1490		1500	·• ,	1510		1520		1530
				•							
	1610	1620		1630		1640		1650		1660	
	CTGCTCCTG										CTAC
	: ::::::										
CCM. AA	CCCCTCCTCC	ጎጥርርሳጥር፣	የሶሶሶ ስት	ጓኔጥዮ ኔተ	200202	ተሞሮልርሰ	~ } } ^ 2 ~ 2	እ ሲር ርር	· ጥር ናር እር	የፖንፈግ	4-4-A4-
GSN: AA	CCGCTCCTG									CACCI	
GSN: AA		CTGCTC:								CACCI	1590
GSN: AA	1540)	1550		1560		1570		1580		
	1540 1670	1680	1550	1690	1560	1700	1570	1710	1580	1720	1590
	1540 1670 AGCATGAAGS	1680 IGCAAGA	1550 AACAA	1690 AGTCC	1560 CTCTT	1700 PATGA	1570 CCTCC1	1710 ACTG	1580 SAGATO	1720 CTCGA	1590
EP0111	1670 AGCATGAAG	1680 TGCAAG	1550 AACAA?	1690 AGTCC	1560 CTCTT	1700 PATGA(1570 CCTCCT	T710	1580	1720 CTCGA	1590
EP0111	1670 AGCATGAAG ::::::::	1680 TGCAAGA :::::: TGCAAGA	1550 AACAA? :::::	1690 AGTCCO	1560 CTCTTT :::::	1700 PATGA(:::::	1570 CCTCC1	T710 PACTGO	1580 SAGATO SAGATO	1720 GCTCGA :::::	1590 ATGCC
EP0111	1670 AGCATGAAG ::::::::	1680 TGCAAG	1550 AACAA? :::::	1690 AGTCCO	1560 CTCTTT :::::	1700 PATGA(:::::	1570 CCTCC1	T710 PACTGO	1580 SAGATO SAGATO	1720 GCTCGA :::::	1590 ATGCC
EP0111	1540 1670 AGCATGAAGT AGCATGAAGT 1600	1680 IGCAAGA :::::: IGCAAGA	1550 AACAAA :::::: AACAAA 1610	1690 AGTCCO ::::: AGTCCO	1560 CTCTTT :::::: CTCTTT 1620	1700 PATGA(:::::	1570 CCTCCT CCTCCT 1630	1710 PACTGO PACTGO	AGATO SAGATO 1640	1720 SCTCGA ::::: SCTCGA	1590 ATGCC
EP0111 GSN: AA	1540 1670 AGCATGAAGT AGCATGAAGT 1600	1680 TGCAAGA TGCAAGA TGCAAGA T740	1550 AACAA? :::::: AACAA? 1610	1690 AGTCC ::::: AGTCC	1560 ETCTTT:::::::::::::::::::::::::::::::::	1700 PATGAG ::::: PATGAG	1570 CCTCCT CCTCCT 1630	1710 PACTGO PACTGO 1770	1580 SAGATO SAGATO 1640	1720 GCTCGA ::::: GCTCGA	TGCC TGCC 1650
EP0111 GSN: AA	1540 1670 AGCATGAAGT AGCATGAAGT 1600 1730 CACCGCCTGC	1680 TGCAAGA TGCAAGA TGCAAGA TT40 CACCACC	AACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1690 AGTCCO :::::: AGTCCO 1750 CAGAGO	1560 CTCTTT CTCTTT 1620 CACCC	1700 FATGAG FATGAG 1760 CAGTCG	1570 CCTCCT :::::: CCTCCT 1630	1710 FACTGO FACTGO 1770 CCCAAC	SAGATO SAGATO SAGATO 1640	1720 SCTCGA ::::: SCTCGA 1780 LAGAGA	TGCC TGCC 1650
EP0111 GSN: AA EP0111	1540 1670 AGCATGAAGT AGCATGAAGT 1600 1730 CACCGCCTGG	1680 TGCAAGA TGCAAGA TGCAAGA TT40 CACCACC	AACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1690 AGTCCC AGTCCC 1750 CAGAGC	1560 CTCTTT CTCTTT 1620 CACCCC	1700 PATGAG STATGAG 1760 CAGTCG	1570 CCTCCT 1630 CTTGTC	1710 PACTGO PACTGO 1770 CCCAAC	JAGATO SAGATO 1640	1720 SCTCGA SCTCGA 1780 LAGAGA	TGCC 1650
EP0111 GSN: AA EP0111	1540 1670 AGCATGAAGT AGCATGAAGT 1600 1730 CACCGCCTGC	1680 TGCAAGA TGCAAGA TGCAAGA T740 CACCACC	AACAAA AACAAA 1610	1690 AGTCCC 1750 CAGAGCCC	1560 CTCTTT 1620 CACCCC	1700 PATGAG PATGAG 1760 CAGTCG	1570 CCTCCT 1630 CTTGTC	T710 FACTGO FACTGO 1770 CCCAAC	1580 EAGATO EAGATO 1640 ETCGAO	1720 GCTCGA ::::: GCTCGA 1780 LAGAGA LAGAGA	1590 ATGCC ATGCC 1650 ACCCT
EP0111 GSN: AA EP0111	1540 1670 AGCATGAAGT AGCATGAAGT 1600 1730 CACCGCCTGC	1680 TGCAAGA TGCAAGA TGCAAGA TT40 CACCACC	AACAAA AACAAA 1610	1690 AGTCCC 1750 CAGAGCCC	1560 CTCTTT 1620 CACCCC	1700 PATGAG PATGAG 1760 CAGTCG	1570 CCTCCT 1630 CTTGTC	T710 FACTGO FACTGO 1770 CCCAAC	1580 EAGATO EAGATO 1640 ETCGAO	1720 GCTCGA ::::: GCTCGA 1780 LAGAGA LAGAGA	1590 ATGCC ATGCC 1650 ACCCT
EP0111 GSN: AA EP0111	1540 1670 AGCATGAAGT 1600 1730 CACCGCCTGC :::::::: CACCGCCTGC	1680 TGCAAGA TGCAAGA TGCAAGA TA40 CACCACA	AACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1690 AGTCCC ::::: 1750 CAGAGC :::::	TCTTT CTCTTT 1620 CACCCC 1680	1700 FATGAG FATGAG 1760 CAGTCG	1570 CCTCCT 1630 CTTGTC 1690	T710 FACTGO FACTGO T770 CCCAAC	ASAGATO SAGATO 1640 STCGAO STCGAO 1700	1720 SCTCGA SCTCGA 1780 CAGAGA CAGAGA	TGCC TGCC TGCC TGCCT TTTO
EP0111 GSN: AA EP0111 GSN: AA	1540 1670 AGCATGAAGT 1600 1730 CACCGCCTGGTCCCCCCCCCCCCCCCCCCCCCCCCCCCC	1680 1680 16CAAGA 11740 1740 CACCACA	AACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1690 AGTCCC ::::: 1750 CAGAGC :::::	TCTTT CTCTTT 1620 CACCCC 1680	1700 FATGAG FATGAG 1760 CAGTCG CAGTCG	1570 CCTCCT 1630 CTTGTC 1690	1710 FACTGO 1770 CCCAAC :::::: CCCAAC	ASAGATO SAGATO 1640 STCGAO STCGAO 1700	1720 SCTCGA 1780 CAGAGA CAGAGA	TGCC TGCC 1650 ACCCT ACCCT T710
EP0111 GSN: AA EP0111 GSN: AA EP0111	1540 1670 AGCATGAAGT 1600 1730 CACCGCCTGC 1660 1790 CCCTCCACCC	1680 IGCAAGA IGCAAGA ITGCAAGA ITGCAAGA ITGCAAGA ITGCAAGA ITGCACACACACACACACACACACACACACACACACACACA	AACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1690 AGTCCC 1750 LAGAGCC 1750 LAGAGCC 1810 GGGTGC	TCTTT CTCTTT 1620 CACCCC 1680	1700 FATGAG 1760 CAGTCG CAGTCG 1820.	1570 CCTCCT 1630 CTTGTC 1690 CGGTTC	1710 FACTGO 1770 CCCAAC 1770 CCCAAC 1830 CTATAT	AGATO SAGATO SAGATO 1640 STCGAC TCAGC	1720 SCTCGA 1780 LAGAGA LAGAGA 1840 ATCTCC	ATGCC LITECC ACCCT ACCCT T710 GAGGC
EP0111 GSN: AA EP0111 GSN: AA EP0111	1540 1670 AGCATGAAGT 1600 1730 CACCGCCTGC 1660 1790 CCCTCCACCC	1680 1GCAAGA 1GCAAGA 1740 CACCACA CACCACACACACACACACACACACACACAC	AACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1690 AGTCCC 1750 CAGAGCC 1750 CAGAGCC 1810 GGGTGC	TCTTT CTCTTT 1620 CACCCC 1680 GAATCC	1700 FATGAG 1760 CAGTCG CAGTCG 1820. GCTCCG	1570 CCTCCT 1630 CTTGTC 1690 CGGTTC	1710 FACTGO 1770 CCCAAC 1770 CCCAAC 1830 CTATAT	AGATO SAGATO SAGATO 1640 STCGAC 1700 CCAGC	1720 SCTCGA 1780 CAGAGA CAGAGA 1840 ATCTCC	ATGCC ATGCC 1650 ACCCT ACCCT T710 GAGGC
EP0111 GSN: AA EP0111 GSN: AA EP0111	1540 1670 AGCATGAAGT 1600 1730 CACCGCCTGC 1660 1790 CCCTCCACCC	1680 1GCAAGA 1GCAAGA 1740 CACCACA CACCACACACACACACACACACACACACAC	AACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1690 AGTCCC 1750 CAGAGCC 1750 CAGAGCC 1810 GGGTGC	TCTTT CTCTTT 1620 CACCCC 1680 GAATCC	1700 FATGAG 1760 CAGTCG CAGTCG 1820. GCTCCG	1570 CCTCCT 1630 CTTGTC 1690 CGGTTC	1710 FACTGO 1770 CCCAAC 1770 CCCAAC 1830 CTATAT	AGATO SAGATO SAGATO 1640 STCGAC 1700 CCAGC	1720 SCTCGA 1780 CAGAGA CAGAGA 1840 ATCTCC	ATGCC ATGCC 1650 ACCCT ACCCT T710 GAGGC
EP0111 GSN: AA EP0111 GSN: AA EP0111	1540 1670 AGCATGAAGT 1600 1730 CACCGCCTGC 1660 1790 CCCTCCACCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	1680 1GCAAGA 1GCAAGA 1740 CACCACA CACCACACACACACACACACACACACACAC	AACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1690 AGTCCC 1750 CAGAGC 1810 GGGTGC	TCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	1700 FATGAG 1760 CAGTCG 1820. GCTCCG GCTCCG	1570 CCTCCT 1630 CTTGTC 1690 CGGTTC	T710 TACTGO TACT	AGATO SAGATO SAGATO 1640 STCGAC 1700 CCAGC	1720 SCTCGA 1780 LAGAGA LAGAGA 1840 ATCTCC	ATGCC ATGCC 1650 ACCCT ACCCT T710 GAGGC
EP0111 GSN: AA EP0111 GSN: AA EP0111	1540 1670 AGCATGAAGT 1600 1730 CACCGCCTGC 1660 1790 CCCTCCACCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	1680 IGCAAGA IGCAAGA ITT40 ITT40 CACCACC ITT40 ITT40 CACCACC	AACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1690 AGTCCC 1750 CAGAGC 1810 GGGTGC	TCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	1700 FATGAG 1760 CAGTCG 1820. GCTCCG GCTCCG	1570 CCTCCT 1630 CTTGTC 1690 CGGTTC	T710 TACTGO TACT	SAGATO SAGATO SAGATO 1640 STCGAO TCAGCA TCAGCA	1720 SCTCGA 1780 LAGAGA LAGAGA 1840 ATCTCC	ATGCC ATGCC 1650 ACCCT ACCCT T710 GAGGC
EP0111 GSN: AA EP0111 GSN: AA EP0111	1540 1670 AGCATGAAGT 1600 1730 CACCGCCTGC 1660 1790 CCCTCCACCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	1680 IGCAAGA IGCAAGA ITT40 ITT40 CACCACC ITT40 ITT40 AGCACCACCACCACCACCACCACCACCACCACCACCACCA	AACAAA 1610 CCCGTC 1670 GGCGGG 1730	1690 AGTCCC 1750 CAGAGC 1810 GGGTGC	TCTTTT 1620 CACCCC 1680 GAATCC 1740	1700 FATGAG 1760 CAGTCG CAGTCG 1820 GCTCCG GCTCCG	1570 CCTCCT 1630 CTTGTC 1690 CGGTTC	T710 FACTGO 1770 CCCAAC CCCAAC 1830 CTATAT	SAGATO SAGATO 1640 STCGAC TOO TCAGCA 1760	1720 SCTCGA 1780 LAGAGA LAGAGA 1840 ATCTCC	ATGCC ATGCC 1650 ACCCT ACCCT T710 GAGGC
EP0111 GSN: AA EP0111 GSN: AA EP0111 GSN: AA	1540 1670 AGCATGAAGT 1600 1730 CACCGCCTGC 1660 1790 CCCTCCACCC 1720 1850	1680 IGCAAGA IGCAAGA ITT40 ITT40 CACCACC ITT40 ITT40 AGCACCACC	AACAAA 1610 CCCGTC 1670 GGCGGG 1730	1690 AGTCCC 1750 LAGAGC 1810 GGGTGC 1817	TCTTT 1620 CACCCC 1680 GAATCC 1740	1700 FATGA(1760 CAGTC(1820. GCTCC(1820.	1570 CCTCCT 1630 CTTGTC 1690 CGGTTC 1750	1710 FACTGO 1770 CCCAAC 1830 CTATAT	SAGATO SAGATO 1640 STCGAC T700 TCAGCA 1760	1720 SCTCGA 1780 1780 CAGAGA LAGAGA 1840 ATCTCG	ATGCC ATGCC ACCCT ACCCT T710 SAGGC 1770
EP0111 GSN: AA EP0111 GSN: AA EP0111 GSN: AA	1540 1670 AGCATGAAGT 1600 1730 CACCGCCTGC 1660 1790 CCCTCCACCC 1720 1850 AGAATCGAGG	1680 IGCAAGA ITGCAAGA ITGCAAGAAGA ITGCAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAA	AACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1690 AGTCCC 1750 CAGAGC :::::: CAGAGC 1810 GGGTGC 1870 AGGCCC	TCTTTT CACCCC CACCCC T680 GAATCC T740	1700 TATGAC TATGAC TATGAC TAGTCC TAGTC TAGTCC TAGTCC TAGTCC TAGTCC TAGTCC TAGTCC TAGTCC TAGTCC TAGTC TAGTCC TAGTC TAGTCC TAGTC TAG	1570 CCTCCT 1630 CTTGTC 1690 CGGTTC 1750 CAGTGT	1710 PACTGO 1770 P	AGATO	1720 CCTCGA ::::: CCTCGA 1780 CAGAGA :::::: CAGAGA ATCTCC	ATGCC 1650 ACCCT CCCT T710 GAGGC 1770 GGTCG
EP0111 GSN: AA EP0111 GSN: AA EP0111 GSN: AA	1540 1670 AGCATGAAG' 1600 1730 CACCGCCTGC 1660 1790 CCCTCCACCC 1720 1850 AGAATCGAGCC 1660	1680 IGCAAGA IGCAAGA ITGCAAGA ITGCAAGAGA ITGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	AACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1690 AGTCCC 1750 CAGAGC 1810 GGGTGC 1870 AGGCCC	TCTTT CACCCC CACCCC T680 GAATCC T740 CCTTTC	1700 TATGAC TATGAC TATGAC TAGTCC TAGTC TAGTCC TAGTC T	1570 CCTCCT 1630 CTTGTC 1690 CGGTTC 1750 CAGTGT	T710 FACTGO FACTGO 1770 CCCAAC FECTATAT	AGATO SAGATO 1640 STCGAC TOAGCA TCAGCA 1760 CAGTAG	1720 CCTCGA 1780 LAGAGA LAGAGA 1840 ATCTCC 1900 LGGAGG	ATGCC ATGCC ATGCC ACCCT ACCCT T710 GAGGC 1770 GGTCG
EP0111 GSN: AA EP0111 GSN: AA EP0111 GSN: AA	1540 1670 AGCATGAAG: :::::::: AGCATGAAG: 1600 1730 CACCGCCTGG: :::::::: CACCGCCTGG 1790 CCCTCCACC: 1720 1850 AGAATCGAGG: AGAATCGAGG	1680 1GCAAGA 1740 1740 CACCACC 1800 AGCAGCC 1860 AGCAGCCC AGTCCGA	AACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1690 AGTCCC 1750 CAGAGC CAGAGC 1810 CGGTGC 1870 AGGCCC	TOTTTO CACCCO CACCO CACCCO CACCCO CACCCO CACCCO CACCCO CACCCO CACCCO	1700 TATGAC TATGAC TATGAC TAGTCC TAGTC TAGTCC TAGTC	1570 CCTCCT 1630 CTTGTC 1690 CGGTTC 1750 CAGTGT	1710 PACTGO 1770 PACTGO 1770 PACCAAO 1830 PACTATATATATATATATATATATATATATATATATATAT	SAGATO SAGATO 1640 STCGAC TOO CAGCA CAGTAT	1720 CCTCGA 1780 LAGAGA LAGAGA 1840 ATCTCC 1900 LGGAGG	ATGCC 1650 ACCCT 1650 ACCCT T710
EP0111 GSN: AA EP0111 GSN: AA EP0111 GSN: AA	1540 1670 AGCATGAAG' 1600 1730 CACCGCCTGC 1660 1790 CCCTCCACCC 1720 1850 AGAATCGAGCC 1660	1680 1GCAAGA 1740 1740 CACCACC 1800 AGCAGCC 1860 AGCAGCCC AGTCCGA	AACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1690 AGTCCC 1750 CAGAGC CAGAGC 1810 CGGTGC 1870 AGGCCC	TCTTT CACCCC CACCCC T680 GAATCC T740 CCTTTC	1700 TATGAC TATGAC TATGAC TAGTCC TAGTC TAGTCC TAGTC	1570 CCTCCT 1630 CTTGTC 1690 CGGTTC 1750 CAGTGT	1710 PACTGO 1770 PACTGO 1770 PACCAAO 1830 PACTATATATATATATATATATATATATATATATATATAT	AGATO SAGATO 1640 STCGAC TOAGCA TCAGCA 1760 CAGTAG	1720 CCTCGA 1780 LAGAGA LAGAGA 1840 ATCTCC 1900 LGGAGG	ATGCC ATGCC ATGCC ACCCT ACCCT T710 GAGGC 1770 GGTCG
EP0111 GSN: AA EP0111 GSN: AA EP0111 GSN: AA	1540 1670 AGCATGAAG: :::::::: AGCATGAAG: 1600 1730 CACCGCCTGG: ::::::: CACCGCCTGG: 1790 CCCTCCACC: :::::::: CCCTCCACC: 1720 1850 AGAATCGAGA: 1780	1680 IGCAAGA IGCAAGA ITGCAAGA ITGCAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAA	ACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1690 AGTCCC 1750 CAGAGC CAGAGC 1810 CGGTGC 1870 AGGCCC	TOTAL CACCO CACCO CACCO TOTAL CACCO TOTAL CACCO TOTAL CACCO TOTAL CACCO TOTAL TO	1700 TATGAC TATGAC TATGAC TAGTCC TAGTC TAGTCC TAGTCC TAGTCC TAGTCC TAGTCC TAGTCC TAGTCC TAGTCC TAGTC TAGTCC TAGTC T	1570 CCTCCT 1630 CTTGTC 1630 CGGTTC 1750 CAGTGT 1810	1710 PACTGO 1770 P	AGATO CAGCA CAGTA CAGTA CAGTA 1820	1720 CTCGA CTCGA 1780 CAGAGA CAGAGA ATCTCC 1900 CGGAGC	ATGCC 1650 ACCCT 1650 ACCCT T710
EP0111 GSN: AA EP0111 GSN: AA EP0111 GSN: AA	1540 1670 AGCATGAAG: :::::::: AGCATGAAG: 1600 1730 CACCGCCTGG: :::::::: CACCGCCTGG: 1790 CCCTCCACC: 1720 1850 AGAATCGAGG: 1780 1910	1680 162AGA 1740 1740 1740 1740 1800 1800 1860 1860 1860 1860 1860 186	AACAAA 1610 CCCGTC 1670 GGCGGGC 1730 AGCAGA 1790	1690 AGTCCC 1750 CAGAGC :::::: CAGAGC 1810 GGGTGC 1870 AGGCCC 1870 AGGCCCC	TOTTO CACCCO CACCCO CACCCO CACCCO TORO CATTO CACCCO TORO	1700 TATGAC 1760 CAGTCC 1820 SCTCCC 1880 GCTCCC 1940	1570 CCTCCT :::::: 1630 CTTGTC :::::: CTGTTGTC 1690 CGGTTC :::::: CGGTTC 1750 CAGTGT	1710 PACTGO 1770 CCCAAC 1770 CCCAAC 1830 CTATAT 1890 FCCTTC	SAGATO SAGATO 1640 STCGAO TOO TCAGCZ TCAGCZ TCAGCZ 1760 CAGTAT	1720 CTCGA ::::: GCTCGA 1780 CAGAGA CAGAGA ATCTCC 1900 CGGAGG CCGGAGG	ATGCC 1650 ACCCT 1710 ACCCT T710
EP0111 GSN: AA EP0111 GSN: AA EP0111 GSN: AA	1540 1670 AGCATGAAG: :::::::: AGCATGAAG: 1600 1730 CACCGCCTGG: :::::::: CACCGCCTGG: 1790 CCCTCCACC: :::::::: CCCTCCACC: 1720 1850 AGAATCGAG: 1780 1910 CGTCCTGACC	1680 16740 1740 1740 1740 1740 1800 1800 1860 1860 1860 1860 1860 1920 1920	AACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1690 AGTCCC 1750 LAGAGC 1810 AGGCCC 1870 AGGCCC 1930 CCTTC	TOTTO CACCO CACCO CACCO CACCO CACCO TOTO	1700 PATGAC PATGAC 1760 PAGTCC PAGTCC 1820 PGCTCCC 1880 PGCTCCCC 1940 PGAGCC	1570 CCTCCT :::::: 1630 CTTGTC :::::: CTGTTGTC 1690 CGGTTC :::::: CGGTTC 1750 CAGTGT	1710 PACTGO 1770 CCCAAC 1770 CCCAAC 1830 CTATAT 1890 FCCTTC	SAGATO SAGATO 1640 STCGAO TOO TCAGCZ TCAGCZ TCAGCZ 1760 CAGTAT	1720 CTCGA ::::: GCTCGA 1780 CAGAGA CAGAGA ATCTCC 1900 CGGAGG CCGGAGG	ATGCC 1650 ACCCT 1710 ACCCT T710
EP0111 GSN: AA EP0111 GSN: AA EP0111 GSN: AA EP0111	1540 1670 AGCATGAAG: :::::::: AGCATGAAG: 1600 1730 CACCGCCTGG: :::::::: CACCGCCTGG: 1790 CCCTCCACC: 1720 1850 AGAATCGAGA: :::::::: AGAATCGAGA: 1780 1910 CGTCCTGACC: ::::::::::::::::::::::::::::::::::	1680 IGCAAGA IGCAAGA ITGCAAGA ITGCAAGA ITGCAACACACACACACACACACACACACACACACACACAC	AACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1690 AGTCCC 1750 LAGAGCC 1810 AGGCCC 1870 AGGCCC 1930 CCTTCC	TOTTO CACCO CACCO CACCO CACCO CACCO TOTO	1700 TATGAC TATGAC TATGAC TAGTCC TAGTC TAGTCC TAGTC TAG	1570 CCTCCT :::::: 1630 CTTGTC :::::: CTGTTGTC 1690 CGGTTC :::::: CGGTTC 1750 CAGTGT	1710 PACTGO 1770 CCCAAC 1770 CCCAAC 1830 CTATAT 1890 FCCTTC	SAGATO SAGATO 1640 STCGAO TOO TCAGCZ TCAGCZ TCAGCZ 1760 CAGTAT	1720 CTCGA ::::: GCTCGA 1780 CAGAGA CAGAGA ATCTCC 1900 CGGAGG CCGGAGG	ATGCC 1650 ACCCT 1710 ACCCT T710
EP0111 GSN: AA EP0111 GSN: AA EP0111 GSN: AA EP0111	1540 1670 AGCATGAAG: :::::::: AGCATGAAG: 1600 1730 CACCGCCTGG: :::::::: CACCGCCTGG: 1790 CCCTCCACC: :::::::: CCCTCCACC: 1720 1850 AGAATCGAG: 1780 1910 CGTCCTGACC	1680 IGCAAGA ITGCAAGA ITGCAAGA ITGCAAGA ITGCACACACACACACACACACACACACACACACACACACA	AACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1690 AGTCCC 1750 LAGAGC 1810 AGGCCC 1870 AGGCCC 1930 CCTTCC	TATTO CACCO CA	1700 PATGAC PATGAC 1760 PAGTCC 1820 SCTCCC 1880 SCTCCCC 1940 PGAGC PGAGC	1570 CCTCCT :::::: 1630 CTTGTC :::::: CTGTTGTC 1690 CGGTTC :::::: CGGTTC 1750 CAGTGT	1710 PACTGO 1770 CCCAAC 1770 CCCAAC 1830 CTATAT 1890 FCCTTC	SAGATO SAGATO 1640 STCGAO TOO TCAGCZ TCAGCZ TCAGCZ 1760 CAGTAT	1720 CTCGA ::::: GCTCGA 1780 CAGAGA CAGAGA ATCTCC 1900 CGGAGG CCGGAGG	ATGCC 1650 ACCCT 1710 ACCCT T710

```
AAB20898 standard; Protein; 620 AA.
XX
     AAB20898;
AC
ХX
                    (first entry)
     05-JAN-2001
DT
XX
     Oryzias lapites oestrogen receptor protein SEQ ID NO:3.
DE
XX
     Oryzias lapites; oestrogen receptor.
KW
XX
     Oryzias lapites.
OS
XX
     JP2000201688-A.
PN
XX
     25-JUL-2000.
PD
XX
                      99JP-0098787.
PF
     06-APR-1999;
\mathbf{x}\mathbf{x}
                      98JP-0319465.
     10-NOV-1998;
```

Land of 500 1 1 VS Sen 7,13 of 50 my 20161

WPI; 2000-567950/53. N-PSDB; AAA92175.

(SUMO) SUMITOMO CHEM CO LTD.

An estrogen receptor gene and its application

Claim 2; Page 13-15; 23pp; Japanese.

The present sequence represents an oestrogen receptor derived from Oryzias lapites. Also described are: (1) a vector comprising the oestrogen receptor gene; (2) a transformant prepared by introducing the oestrogen receptor gene or vector from (1) into a host cell; (3) a method for the preparation of an oestrogen receptor comprising culturing the transformant from (2) to produce the oestrogen receptor; and (4) a method for the evaluation of oestrogen receptor-activating ability of a chemical substance in which the chemical substance is reacted with a transformant prepared by introducing a reporter gene connected downstream of a transcription controlling region containing an oestrogen response sequence and the above oestrogen receptor gene to an oestrogen-nonendogenous host cell. The transformant can be used for the evaluation of oestrogen receptor-activating ability of a chemical substance.

620 AA; SO 39 A; 44 R; 10 N; 29 D; 0 B; 18 C; 27 Q; 31 E; 0 Z; 60 G; 19 H; SQ 24 I; 67 L; 25 K; 20 M; 15 E; 43 P; 67 S; 30 T; 4 W; 20 Y; 28 V; SQ mskrqssvqi rqlfgpalrs rispassele tlspprlspr dplgdmypee srgsggvaav dllegtydya apnpattply sqsstgyysa pletngppse gslqslqsqp tsplvfvpss prlspfmhpp shhylettst pvyrsshqga sredqcgsre dtcslgelga gagaggfema kdtrfcavcs dyasgyhygv wscegckaff krsigghndy mcpatngcti drnrrkgcga crlrkcyevg mmkggvrkdr irilrrdkrr tgvgdgdkvv kgqehktvhy dgrkrsstgg gggggggrls vtsippeqvl 111qqaeppi learqhlarp ytevtmmtll tsmadkelvh miawakklpg flqlslhdqv lllesswlev lmigliwrsi hcpgklifaq dlildrnegd cvegmteifd mllatasrfr vlklkpeefv clkaiillns gefsfctgtm eplhnsaavqsmldtitdal ihyisqsgyl aqeqarrqaq pllllshirh msnkgmehly smkcknkvpl ydlllemlda brlhhpvrap gslsqvdrdp pstssggggi apgsisasrg riespsrgpf apsvlqyggs rpdctpalqd

//

PR XX

PA $\mathbf{x}\mathbf{x}$

DR

DR XX

PT $\mathbf{x}\mathbf{x}$

PS XX

CC

CC CC

XX

```
Smith-Waterman score: 4198; 99.516% identity (99.516% ungapped) in 620 aa overlap
(75-1934:1-620)
                               180
               120
                       150
        90 .
EP0111 MSKRQSSVQIRQLFGPALRSRISPASSELETLEPPRLSPRDFLGDMYPEESRGSGGVAAV
     GSP: AA MSKRQSSVQIRQLFGPALRSRISPASSELETLSPPRLSPRDPLGDMYPEESRGSGGVAAV
                                   40
                                           50
                                                   60
                           30
                   20
                               360
                                       390
                                               420
                       330
        270
               300
EP0111 DFLEGTYDYAAPNPATTPLYSQSSTGYYSAPLETNGPPSEGSLQSLGSGPTSPLVFVPSS
     GSP: AA DLLEGTYDYAAPNPATTPLYSQSSTGYYSAPLETNGPPSEGSLQSLGSGPTSPLVFVPSS
                                   100
                                          110
                                                   120
                  80
                           90
                                       570
                                               600...
               480
                       510 ·
                               540 --
EP0111 PRLSPFMHPPSHHYLETTSTPVYRSSHQGASREDQCGSREDTCSLGELGAGAGAGGFEMA
     GSP: AA PRLSPFMHPPSHHYLETTSTPVYRSSHQGASREDQCGSREDTCSLGELGAGAGAGGFEMA
                                          170
                           150
                                   160
           130
                   140
                                       750
                       690_
                               7.20_
                660
        630.
EP0111 KDTRFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYMCPATNQCTIDRNRRKSCQA
     GSP: AA KDTRFCAVCSDYASGYHYGVWSCEGCKAFFERSIQCHNDYMCPATNQCTIDRNRRKGCQA
                                                   240
                                   220
                                          230
                           210
                   200
           190
                              900__
                                       930
               840
                       870....
        810
EP0111 CRLRKCYEVGMMKGGVRKDRIRILRRDKRRTGVGDGDKVVKGQEHKTVHYDGRKRSSTGG
      .................
GSP: AA CRLRKCYEVGMMKGGVRKDRIRILRRDKRRTGVGDGDKVVKGQEHKTVHYDGRKRSSTGG
                                           290
                                                   300
                                   280
                           270
           250
                   260
                       1050
                               1080
                                       1110
                                               1140
               1020
        990
EP0111 GGGGGGGRLSVTSIPPEQVLLLLQGAEPPILCSRQKLSRPYTEVTMMTLLTSMADKELVH
      ...............
GSP: AA GGGGGGGRLSVTSIPPEQVLULLQGAEPPILCSRQKLSRPYTEVTMMTLLTSMADKELVH
                                                   360
                                   340
                                           350
                   320
                           330
           310
                                       1290
               1200
                      1230
                              1260
       1170
EP0111 MIAWAKKLPGFLQLSLHDQVLLLESSWLEVLMIGLIWRSIHCPGKLIFAQDLILDRNEGD
      GSP: AA MIAWAKKLPGFLQLSLHDQVLLLESSWLEVLMIGLIWRSIHCPGKLIFAQDLILDRNEGD
                                                   420
                                           410
                           390
                                   400
           370
                   380
                                               1500
                               1440
                                       1470...
               1380
                       1410
EPO111 CVEGMTEIFDMLLATASRFRVLKLKPEEFVCLKAIILLNSGAFSFCTGTMEPLHNSAAVQ
      GSP: AA CVEGMTEIFDMLLATASRFRVLKLKPEEFVCLKAIILLNSGAFSFCTGTMEPLHNSAAVQ
                                                   480
                                   460
                                           470
                   440
                           450
           430
                       1590
                               1620___
                                       16.50 ...
               1560
       1530...
EP0111 SMLDTITDALIHYISQSGYLAQEQARRQAQLLLLLSHIRHMSNKGMEHLYSMKCKNKVPL
      GSP: AA SMLDTITDALIHYISQSGYLAQEQARRQAQPLLLLSHIRHMSNKGMEHLYSMKCKNKVPL
                                           530
                                                  540
                           510
                                   520
                   500
           490
                                       1830
               1740
                       1770
                               1800
EPO111 YDLLLEMLDAHRLHHPVRAPQSLSQVDRDPPSTSSGGGGIAPGSISASRGRIESPSRGPF
      GSP: AA YDLLLEMLDAHRLHHPVRAPQSLSQVDRDPPSTSSGGGGIAPGSISASRGRIESPSRGPF
                                                   600
                           570
                                           590
           550
                                   580
                   560
```

>>GSP:AAB20898 Oryzias lapites Qestrogen recept (620 aa)

initn: 4198 init1: 4198 opt: 4198 Z-score: 3093.7 bits: 584.7 E(): 2.3e-164

AAA92174;

ID

ХX

AC

XX

DT XX

DE XX

KW XX

os

XX PN

XX PD

XX PF

XX

PR XX

PA XX

DR DR

XX

PT XX

PS XX

CC

SQ

XP-002181484

AAA92174;

05-JAN-2001 (first entry)

Oryzias lapites oestrogen receptor encoding DNA SEQ ID NO: 2.

P. 05-01-2001

1-7

Oryzias lapites; oestrogen receptor; ds.

Oryzias lapites.

JP2000201688-A.

25-JUL-2000.

06-APR-1999; 99JP-0098787.

10-NOV-1998; 98JP-0319465.

(SUMO) SUMITOMO CHEM CO LTD.

WPI; 2000-567950/53. P-PSDB; AAB20897.

An estrogen receptor gene and its application

Claim 3; Page 11-13; 23pp; Japanese.

The present sequence encodes an oestrogen receptor derived from Oryzias lapites. Also described are: (1) a vector comprising the oestrogen receptor gene; (2) a transformant prepared by introducing the oestrogen receptor gene or vector from (1) into a host cell; (3) a method for the preparation of an oestrogen receptor comprising culturing the transformant from (2) to produce the oestrogen receptor; and (4) a method for the evaluation of oestrogen receptor-activating ability of a chemical substance in which the chemical substance is reacted with a transformant prepared by introducing a reporter gene connected downstream of a transcription controlling region containing an oestrogen response sequence and the above oestrogen receptor gene to an oestrogen-nonendogenous host cell. The transformant can be used for the evaluation of oestrogen receptor-activating ability of a chemical substance.

Sequence 1728 BP; 378 A; 514 C; 497 G; 339 T; 0 other; atgtaccetg aagagageeg gggttetgga ggggtggetg etgtggaeet tttggaaggg 60 acgtacgact atgccgccc caaccttgcc acgactcccc tttacagcca gtccagcacc 120 ggctactact ctgctcccct ggaaacaaac ggacccccct cagaaggcag tctgcagtcc 180 ctgggcagtg ggccgacgag ccctctggtg tttgtgccct ccagccccag actcagtccc 240 300 tttatgcatc cacccagcca ccactatctg gaaaccactt ccacgcccgt ttacagatcc agccaccagg gagcctccag ggaggaccag tgcggctccc gggaggacac gtgcagcctg 360 ggggagttag gcgccggagc cggggctggg gggtttgaga tggccaaaga cacgcgtttc 420 tgcgccgtgt gcagcgacta cgcctctggg taccactatg gggtgtggtc ttgtgagggc 480 tgcaaggcct tcttcaagag gagcatccag ggtcacaatg actatatgtg cccagcgacc 540 600 aatcagtgca ctattgacag aaatcgaagg aagggctgtc aggcttgtcg tcttaggaag tgttacgaag tgggaatgat gaaaggcggt gtgcgcaagg accgcattcg cattttacgg 660 720 cgtgacaaac ggcggacagg cgttggtgat ggagacaagg ttgtaaaggg tcaggagcat aaaacggtgc attatgatgg aaggaaacgc agcagcacag gaggaggagg aggaggagga 780 ggaggaagac tgtctgtgac cagcatacct cctgagcagg tgctgctcct ccttcagggc 840 900 gccgagcccc cgatactctg ctcgcgtcag aagttgagcc gaccgtacac cgaggtcacc atgatgaccc tgctcaccag catggcagac aaggagctgg tccacatgat cgcctgggcc 960 1020 aagaagctcc caggttttct gcagctgtcc ctgcacgatc aggtgctgct gctggagagc tcgtggctgg aggtgctcat gatcggcctc atttggaggt ccatccactg tcccgggaag 1080 1140 ctcatctttg cacaagacct catcctggac aggaatgagg gagactgcgt ggaaggcatg

acqqaqatct	tcgacatgct	gctggccact	gcttcccgct	tccgtgtgct	caaactcaaa	1200
cctgaggaat	tcgtctgcct	caaagctatt	attttactca	actccggtgc	tttttctttc	1260
tgcaccggca	ccatggagcc	acttcacaac	agcgcggcgg	ttcagagcat	gċtggacacc	1320.
atcacagacg	cactcattca	ttacatcagt	cagtcgggtt	acttggccca	ggagcaggcg	1380
agacggcagg	cccagccgct	cctgctgctc	tcccacatca	ggcacatgag	caacaaaggc ·	1440
atggagcacc	tctacagcat	gaagtgcaag	aacaaagtcc	ctctttatga	cctcctactg	15 00
gagatoctco	atacccacca	cctgcaccac	cccgtcagag.	cccccagtc	cttgtcccaa.	15.60.
gtcgacagag	accctccctc	caccagcagc	ggcgggggtg	gaatcgctcc	cggttctata	1620
tcagcatctc	gaggcagaat	cgagagtccg	agcagaggcc	cctttgctcc	cagtgtcctt	1680
			ccggcccttc			1728

```
>>GSN:AAA92174 Oryzias lapites oestrogen recept (1728 nt)
initn: 8586 init1: 8586 opt: 8586 Z-score: 8271.0 bits: 1544.2 E():
99.653% identity (99.653% ungapped) in 1728 nt overlap (211-1938:1-1728)
                200
                       210
                              220
                                     230
EP0111 CGCCTCTCGCCCCGTGACCCCCTCGGTGACATGTACCCTGAAGAGAGCCCGGGGTTCTGGA
                         ATGTACCCTGAAGAGAGCCGGGGTTCTGGA
GSN:AA
                              280
                                     290
                                            300
                       270
EP0111 GGGGTGGCTGTGGACTTTTTGGAAGGGACGTACGACTATGCCGCCCCCAACCCTGCC
     GSN: AA GGGGTGGCTGTGGACCTTTTGGAAGGGACGTACGACTATGCCGCCCCCAACCCTGCC
                 50
                        60
                               70
                                      80
          40
                              340 -
                       330
                320
         310
110
                       120
                              130
                                     140
         100
                      390
                380
                              400 .
                                     410
         370
EP0111 GGACCCCCTCAGAAGGCAGTCTGCAGTCCCTGGGCAGTGGGCCGACGAGCCCTCTGGTG
     GSN: AA GGACCCCCTCAGAAGGCAGTCTGCAGTCCCTGGGCAGTGGGCCGACGAGCCCTCTGGTG
                                     200
                                            210
                       180
                              190
         160
                170
                440
                       450
                              460
                                     470 '
         430
GSN: AA TTTGTGCCCTCCAGCCCCAGACTCAGTCCCTTTATGCATCCACCCAGCCACCACTATCTG
         220
                230
                       240
                              250
                              520
         490
                500
                       510
EP0111 GAAACCACTTCCACGCCCGTTTACAGATCCAGCCACCAGGGAGCCTCCAGGGAGGACCAG
     GSN: AA GAAACCACTTCCACGCCCGTTTACAGATCCAGCCACCAGGGAGCCTCCAGGGAGGACCAG
                                            330
                290
                       300
                              310
         280
                       570
                              580
                                     590
         550
                560
EP0111 TGCGGCTCCCGGGAGGACACGTGCAGCCTGGGGGAGTTAGGCGCCCGGAGCCGGGGCTGGG
     GSN: AA TGCGGCTCCCGGGAGACACGTGCAGCCTGGGGGAGTTAGGCGCCGGAGCCGGGGCTGGG
                                            390
                350
                       360
                                     380
         340
                                            660
                620-
                       630
                              640
                                     650 -
         610
EP0111 GGGTTTGAGATGGCCAAAGACACGCGTTTCTGCGCCGTGTGCAGCGACTACGCCTCTGGG
     .................
GSN: AA GGGTTTGAGATGGCCAAAGACACGCGTTTCTGCGCCGTGTGCAGCGACTACGCCTCTGGG
                                            450
                410
                       420
         400
                              700 710
                       69.0
         670
                680
EP0111 TACCACTATGGGGTGTGGTCTTGTGAGGGCTGCAAGGCCTTCTTCAAGAGGAGCATCCAG
     GSN: AA TACCACTATGGGGTGTGGTCTTGTGAGGGCTGCAAGGCCTTCTTCAAGAGGAGCATCCAG
                       480
                                            "510"
                              490
          460
                470
                       750
                              760
          730
                740
EP0111 GGTCACAATGACTATATGTGCCCAGCGACCAATCAGTGCACTATTGACAGAAATCGGAGG
     GSN: AA GGTCACAATGACTATATGTGCCCAGCGACCAATCAGTGCACTATTGACAGAAATCGAAGG
                                     560
                                             570
          520
                              550
                530
                       540
```

	*						
		790 AAGAGCTGCCAGG	800	810	820	830 Gargargaái	840 • CCCCCT:
	EP0111	AAGAGCTGCCAGGC	TTGTCGTCT	PAGGAAGIGIT	ACGAAGIGG	GAATGATGAA	
	001.33	AAGGGCTGTCAGG	これないという	PAGGAAGTGTT	TACGAAGTGG	GAATGATGAA	AGGCGGT
٠	GSN: AA	580	590	600	610	620	630
		500					
•		850	860	870	880	890	900
	EP0111	GTGCGCAAGGACCC	CATTCGCATT	TTACGCCGT	GACAAACGGC	GGACAGGCGT'	rggtg at
		::::::::::::::		::::::::::	::::::::	::::::::	::::::
	GSN:AA	GTGCGCAAGGACC	CATTCGCATT	PTTACGGCGT	Gac abaegge	GGACAGGCGT:	rggtgat
		640	650	660	670	680	690
		96 - 5					
		910	920	930	940	950	3 6 0.
	EP0111	GGAGACAAGGTTGT	PAAAGGGTCA	GAGCATAAA	ACGGTGCATT	ATGATGGAAG	GAAACGC ·
		:::::::::::::			: : : : : : : : : : : : : : : : : : :	::::::::::	::::::
	GSN: AA	GGAGACAAGGTTGT				ATGATGGAAG 740	
		700	710	720 ···	730	740	750
		970	980	990	1000	1010	102 0
	0111	AGCAGCACAGGAG					
	ELOTIT	AGCAGCACAGGAGG		······	·····	::::::::::	::::::
	CCNLAA	AGCAGCACAGGAG	aggag gag gi	AGGAGGAGGA(GAAGACTOT	CTGTGACCAG	CAT AC CT
	GSW:AA	760	770	7 80	790	800	810
		1030		1050		1070	1080
	EP0111	CCTGAGCAGGTGCT	rgctcctcc t :	rcagggcgcc	BAGCCCCCGA	PACTCTGCTC	GCGTCAG
		:::::::::::::	:::::::::::			Trittiir::T	
	GSN: AA	CCTGAGCAGGTGC	rgctectcct:	rca gggggccc			
		820	830	840	850 _.	·860	870
	•	•					7740
	•	1090	1100	1110	1120	1130	1140
	EP0111	AAGTTGAGCCGAC					
		AAGTTGAGCCGAC					
	GSN: AA	AAGTTGAGCCGACG	890	900	910	920	930
		860	890	300	320	, , ,	
		1150	1160	1170	1180	1190	1200
٠.	EP0111	AAGGAGCTGGTCC			AAGCTCCCAG	GTTTTCTGCA	GCTGTCC
		:::::::::::::		:::::::::	:::::::::	*****	::::::
	GSN:AA	AAGGAGCTGGTCC	ACATGATCGC	CTGGGCCAAG	AA'GCTCCCA	GITITCTGCA	GCIGICC .
		940	950	9 60	370	980	990
					32.22		
			1220		1240	1250	1260
	EP0111	CTGCACGATCAGG	rgctgctget	OCACACE TCG	regero gase	PGC PCATGAT	COSCETE
		TTT:::::::::::::::::::::::::::::::::::	::::::::::::::::::::::::::::::::::::::	#1111111111111 CCN CN CCTCC	7:: 77 7:::: TCC CT CCAC	መሮርም ሮ አ ጥር አ ጥ	CCCCCCC
	GSN: AA	1000	1010		1030		1050
	•	1000	1010	1020		-,	
		1270	1280	12 90	1300	1310	1320
	EP0111	ATTTGGAGGTCCA'	TCCACTGTCE	eg caac ete	atcttt c ca(A agae etc at	CETGGAC
			:::::::::	::::::::::	::::::::::		*******
	GSN:AA	ATTTGGAGGTCCA	TCCACTGTCC	CGGGAAGCTC	ATCTTTGCA		CCTGGAC
		1060	1070	1080	109 0	1100	1110
		1330-	1340 _	1350		1370	1380
	EP0111	AGGAATGAGGGAG.	actgcgtg ga	ассе ат са сс	C acat ette(CACATCETCET	GGCCAC1
		. : : : : : : : : : : : : : : : : : : :	::::::::::				CCCCACE
	GSN:AA	AGGAATGAGGGAG.			GAGATETTE	1160	1170
		1120	1130	1140	1120	1100	11/0
		1390	1400	1410	1420	1430	1440
	ED0111	GCTTCCCGCTTCC					
	PEATIT	FILLIFIE					
	GSM - AA	GCTTCCCGCTTCC	GTGTGCTCAA	ACTCAAACCT	GAGGAATTC	STCTGCCTCAL	AGCTATT
	CDIV.AA	1180	1190	1200	1210	7220	1230

	1450	1460	1470	1480	1490	1500
EP0111	ATTTTACTCAAC	TCCGGTGCTTT	TTCTTTCTGC	ACC CCACC	ATGGAGCGAC1	TEACAAC
GSN: AA	::::::::::::::::::::::::::::::::::::::					
	1240	. 1250	1 260	1270	1280	1290
	1510	1520	1530	1540	1550	15 60
EP0111	AGCGCGGCGGTT					
	:::::::::::::::::::::::::::::::::::::::					
GSN:AA	AGCGCGGCGGTT					
	1300	1310	1320	1330	1340	1350
•	1570	1580	1590	1600	1610	1620
EP0111	CAGTCGGGTTAC	TTGGCCCAGGA	GCAGGCGAGA	CGGCAGGCCC	AGCTGCTCCT	GCTGETE
GSN:AA	CAGTCGGGTTAC		GCAGGCGAGA	CGGCAGGCCC	AGCCGETCCI	GETGETC
	1360	1370	1380	1390	1400	1410
	1630	1640	1650	1660	1670	1680
EP0111	TCCCACATCAGG					
	:::::::::::::::::::::::::::::::::::::::					
GSN:AA	TCCCACATCAGG	*				
	1420	1430	1440	1450	1460	···1470
	1690	1700	1710	1720	1730	1740
EP0111	AACAAAGTCCCT	СТТТАТСАССТ	CC TAC T CCAC	atc cteg atc	EECAECGCCT	GEACEAC
	-::::::::::::::::::::::::::::::::::::::					
GSN: AA	AACAAAGTCCCT		CCTACTCCAG	atgetegatg	CCCACCGCCT	GCACCAC
	1480	1490	1500	1510	1520	1530
	1750	1760	1770	1780	1790	1800
EP0111	CCCGTCAGAGCA	CCCCAGTCCTT	STECCAAGTE	Cacacacac c	CTECCTECAC	CAGCAGC
	***********	:::::::	:: :::::	*********	*********	::-:-:
GSN: AA	CCCGTCAGAGCC			GACAGAGACC	CTCCCTCCAC	CAGCAGC
	1540	1550	1560	1570	15 80	1590
	1810	1820	1830	1840	1850	1860
EP0111	GGCGGGGGTGGA		the state of the s			
CCM. NA	GGCGGGGGTGGA					
GDIV.AA	1600	1610	1620	1630	1640	1650
•		1010	1020		1040	1630
•	1870	1880	1890	1900	1910	1920
EP0111	AGCAGAGGCCCCT					
GSN:AA	AGCAGAGGCCCCT					
y . •	1660	167 0	1680	1690	1700	1710
	1930	1940	1950	1960	1970	1980
EP0111	CCGGCCCTTCAAC	SACTGAGCACAC	CAGTCCAAGG	CCCTTTTTTT	GTGGCTCAAG	GGTTCAG
		:::::		•	*	
GSN:AA	CCGGCCCTTCAAC	SACTGA				
	1720					
	•	,			•	

```
AAB20897 standard; Protein; 575 AA.
ΙĎ
ХX
     AAB20897;
AC
\mathbf{x}\mathbf{x}
      05-JAN-2001
                     (first entry)
\mathbf{DT}
XX
     Oryzias lapites oestrogen receptor protein SEQ ID NO:1.
DE
XX
      Oryzias lapites; oestrogen receptor.
KW
XX
      Oryzias lapites.
os
XX
      JP2000201688-A.
PN
XX
      25-JUL-2000.
PD
ХX
                       99JP-0098787.
      06-APR-1999;
PF
XX
                      98JP-0319465.
      10-NOV-1998;
PR
XX
      (SUMO ) SUMITOMO CHEM CO LTD.
PA
\mathbf{x}\mathbf{x}
      WPI; 2000-567950/53.
DR
      N-PSDB; AAA92174.
DR
XX
      An estrogen receptor gene and its application
PT
XX
      Claim 1; Page 9-10; 23pp; Japanese.
PS
```

The present sequence represents an oestrogen receptor derived from Oryzias lapites. Also described are: (1) a vector comprising the oestrogen receptor gene; (2) a transformant prepared by introducing the oestrogen receptor gene or vector from (1) into a host cell; (3) a method for the preparation of an oestrogen receptor comprising culturing the transformant from (2) to produce the oestrogen receptor; and (4) a method for the evaluation of oestrogen receptor-activating ability of a chemical substance in which the chemical substance is reacted with a transformant prepared by introducing a reporter gene connected downstream of a transcription controlling region containing an oestrogen response sequence and the above oestrogen receptor gene to an oestrogen-nonendogenous host cell. The transformant can be used for the evaluation of oestrogen receptor-activating ability of a chemical substance.

```
\mathbf{x}\mathbf{x}
     Sequence 575 AA;
SQ
     37 A; 38 R; 10 N; 27 D; 0 B; 18 C; 24 Q; 29 E; 0 Z; 58 G; 19 H;
SO
     22 I; 61 L; 24 K; 19 M; 14 F; 37 P; 58 S; 29 T; 4 W; 20 Y; 27 V;
SQ
SO
     0 Others;
     mypeesrgsg gvaavdlleg tydyaapnpa ttplysgsst gyysapletn gppsegslgs
     lgsqptsplv fvpssprlsp fmhppshhyl ettstpvyrs shqgasredq cqsredtcsl
     gelgagagag gfemakdtrf cavcsdyasg yhygvwsceg ckaffkrsiq ghndymcpat
     ngctidrnrr kgcqacrlrk cyeygmmkgg yrkdririlr rdkrrtgygd gdkyvkggeh
     ktvhydgrkr sstggggggg ggrlsvtsip peqvllllqg aeppilcsrq klsrpytevt
     mmtlltsmad kelvhmiawa kklpgflqls lhdqvllles swlevlmigl iwrsihcpgk
     lifaqdlild rnegdcvegm teifdmllat asrfrvlklk peefvclkai illnsgafsf
     ctgtmeplhn saavqsmldt itdalihyis qsgylaqeqa rrqaqpllll shirhmsnkq
     mehlysmkck nkvplydlll emldahrlhh pvrapqslsq vdrdppstss ggggiapgsi
     sasrgriesp srgpfapsvl qyggsrpdct palqd
```

//

XX

CC

```
initn: 3905 init1: 3905 opt: 3905 Z-score: 2879.0 bits: 544.9 E(): 2.1e-152
Smith-Waterman score: 3905; 99.478% identity (99.478% ungapped) in 575 aa overlap
(210-1934:1-575)
    210
                    270
                            300
                                    330
                                            360
EP0111 MYPEESRGSGGVAAVDFLEGTYDYAAPNPATTPLYSQSSTGYYSAPLETNGPPSEGSLOS
      GSP: AA MYPEESRGSGGVAAVDLLEGTYDYAAPNPATTPLYSQSSTGYYSAPLETNGPPSEGSLOS
                    20
                            30
                                    40
                                            50
                                                    60
    390
            420
                    450
                            480
                                    510
                                            540
EP0111 LGSGPTSPLVFVPSSPRLSPFMHPPSHHYLETTSTPVYRSSHQGASREDQCGSREDTCSL
      GSP: AA LGSGPTSPLVFVPSSPRLSPFMHPPSHHYLETTSTPVYRSSHQGASREDQCGSREDTCSL
            70
                    80
                            90
                                   100
                                           110
            600
                    630
                                    690
                            660
EP0111 GELGAGAGAGGFEMAKDTRFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYMCPAT
      GSP: AA GELGAGAGAGGFEMAKDTRFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYMCPAT
           130
                   140
                           150
                                   160
                                           170
            780
    750
                    810
                            840
                                   870
                                           900
EP0111 NQCTIDRNRRKSCQACRLRKCYEVGMMKGGVRKDRIRILRRDKRRTGVGDGDKVVKGQEH
      GSP: AA NOCTIDRNRRKGCQACRLRKCYEVGMMKGGVRKDRIRILRRDKRRTGVGDGDKVVKGOEH
           190
                   200
                           210
                                   220
                                          230
            960
                    990
                           1020
                                   1050
    930
                                          1080
EP0111 KTVHYDGRKRSSTGGGGGGGGRLSVTSIPPEQVLLLLQGAEPPILCSRQKLSRPYTEVT
      GSP: AA KTVHYDGRKRSSTGGGGGGGGGRLSVTSIPPEQVLLLLQGAEPPILCSRQKLSRPYTEVT
           250
                   260
                           270
                                   280
                                          290
                                                  300
   1110
           1140
                   1170
                           1200
                                   1230
                                          1260
EP0111 MMTLLTSMADKELVHMIAWAKKLPGFLQLSLHDQVLLLESSWLEVLMIGLIWRSIHCPGK
     GSP: AA MMTLLTSMADKELVHMIAWAKKLPGFLQLSLHDQVLLLESSWLEVLMIGLIWRSIHCPGK
                   320
                           330
   1290
           1320
                   1350
                           1380
                                   1410
                                          1440
EP0111 LIFAQDLILDRNEGDCVEGMTEIFDMLLATASRFRVLKLKPEEFVCLKAIILLNSGAFSF
     GSP: AA LIFAQDLILDRNEGDCVEGMTEIFDMLLATASRFRVLKLKPEEFVCLKAIILLNSGAFSF
           370
                   380
                           390
                                  400
                                          410
           1500
                   1530
                           1560
                                  1590
EP0111 CTGTMEPLHNSAAVQSMLDTITDALIHYISQSGYLAQEQARRQAQLLLLLSHIRHMSNKG
      GSP: AA CTGTMEPLHNSAAVQSMLDTITDALIHYISQSGYLAQEQARRQAQPLLLLSHIRHMSNKG
           430
                   440
                           450
                                  460
                                          470
                                                  480
   1650
           1680
                   1710
                           1740
                                  1770
                                          1800
EP0111 MEHLYSMKCKNKVPLYDLLLEMLDAHRLHHPVRAPQSLSQVDRDPPSTSSGGGGIAPGSI
     GSP: AA MEHLYSMKCKNKVPLYDLLLEMLDAHRLHHPVRAPQSLSQVDRDPPSTSSGGGGIAPGSI
                                  520
           490
                  500
                          510
                                          530
                                                  540
   1830
           1860
                  1890
                           1920
EP0111 SASRGRIESPSRGPFAPSVLQYGGSRPDCTPALQD
     GSP: AA SASRGRIESPSRGPFAPSVLQYGGSRPDCTPALQD
```

>>GSP:AAB20897 Oryzias lapites oestrogen recept (575 aa)